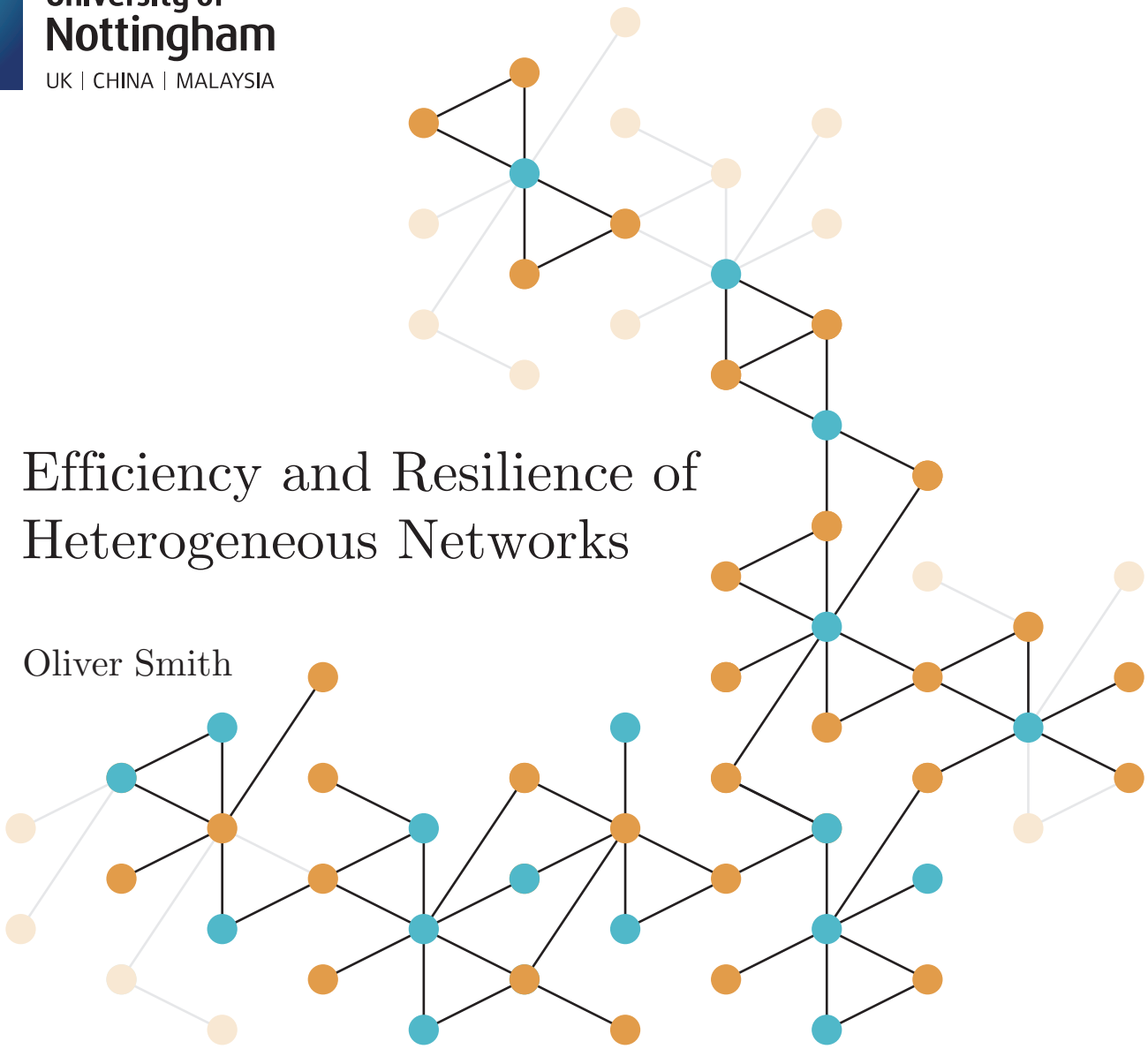
University of Nottingham
UK | CHINA | MALAYSIA



# Efficiency and Resilience of Heterogeneous Networks

Oliver Smith

A thesis presented for the degree of
Doctor of Philosophy

Supervised by:
Reuben O'Dea, Keith Hopcraft, John Crowe and Etienne Farcot

School of Mathematical Sciences,
University of Nottingham, UK
June 2021

# Abstract

Many systems such as traffic or electrical flow can be described as flows following paths of least resistance in networks. The efficiency and resilience of these networks define the system's ability to function effectively. Research into network efficiency and resilience often focuses on the role of network topology, with the aim of uncovering optimal network structures that boost system performance. However, little attention has been paid to the role of node behaviour. This thesis bridges that gap by analysing the efficiency and resilience of networks whose nodes have heterogeneous behaviour. The nodes may variably be sources or sinks of the flow. The nodes may also be equipped with the ability to adjust their behaviour in response to the state of the network. The efficiency and resilience of networks are evaluated as a function of their composition of node types and behaviours. The primary motivation for this is the proliferation of renewable sources of electrical power in energy grids. The resulting electrical networks have highly dynamic and heterogeneous nodes. This thesis provides a framework in which to analyse the behaviour of these systems.

A variety of mathematical methods are utilised throughout this thesis. The efficiency of network flows is analysed using a measurement from game theory called the Price of Anarchy, from which an equivalency between least-resistance network flows and Nash equilibria is also identified. The average variation of efficiency with node composition is found to be approximately invariant across different network structures. The highest inefficiencies are found to always occur when there are an equal number of source and sink nodes.

Resilience is investigated using models of cascading network failures. Both a steady state and a dynamical model are employed. Analytical results for cascades on simple lattices are derived, while for complex networks it is shown that resilience can often be improved by increasing the numbers of source and sink nodes. This analysis is employed on a test case of electrical networks, constructed using real household power consumption and photo-voltaic generation data. The impact of the dynamic variability of these data-driven networks on resilience is analysed. Lowest resiliences are found during times when high numbers of photo-voltaic source nodes are active.

# Acknowledgements

Studying for a Ph.D. and writing this thesis would not have been possible without the host of people who have supported me. First and foremost I must thank my parents, Linda and Paul, who have shown me boundless patience and encouragement over the years. They've been truly selfless and I can't express my appreciation enough. My partner Sam has also shown me nothing but love and support: enduring my worst bouts of grumpiness and, against the odds, surviving three years of my cookery. I'm also of course indebted to my quartet of supervisors. They've given me the opportunity to work in a fascinating field, and their patience and guidance has kept me on track. Keith and John even went beyond the call of duty and remained involved post-retirement, attaining a sort of cyber Obi-Wan Kenobi like status. Together, my supervisors have kept the show on the road even despite a pretty disagreeable pandemic.

The research presented within this thesis was conducted as part of the University of Nottingham's "Mathematics and Analytics for a Sustainable Society" (MASS) project. I want to thank the Leverhulme Trust for funding MASS and providing me with the means not only to conduct my research, but also to attend conferences and work within a fun and stimulating research environment. I want to especially thank Markus Owen, the organiser of MASS, who has consistently fought for everyone within the group. He's been incredibly generous throughout the pandemic, looking after everyone and ensuring they have the support they need.

# Table of Contents

# Summary

This summary is intended to serve as a high level guide or map to the thesis. The reader may wish to refer back here periodically in order to contextualise a particular chapter in the overall theme of the work.

The objective of this thesis is to analyse the efficiency and resilience of flows on networks that possess nodes with heterogeneous behaviour. This will uncover the dependence of network performance on node properties. The primary application of this work is to power grids containing high volumes of variable, low output renewable power. A diverse set of methods will be employed throughout the thesis. Chapter 1 introduces the main themes and reviews the literature. Each section in the chapter will briefly review a particular subject that will be utilised later on.

Chapter 2 is concerned with evaluating the efficiency of flows following paths of least resistance in networks. The chapter exploits the equivalence between these flows and Nash equilibria. This enables the Price of Anarchy, a measurement from game theory, to be used as a computationally efficient means of computing (in)efficiency. The relationship between efficiency and network redundancy is then discussed. The chapter then moves on to compute the efficiency of complex networks as a function of the proportions of flow source and sink nodes. These results are then interpreted for the design of renewable electrical grids. The research in chapter 2 is published in Smith et al. (2019).

Chapter 3 assesses the resilience of electrical grids to cascading failures. The chapter begins by introducing the steady state cascade algorithm that will be used to compute resilience. The resilience of simple lattice networks is then investigated, and analytic results derived. These results will provide intuition for the case of large, complex networks which will be investigated using numerical simulations. The resilience of these complex networks will be computed as a function of their proportions of flow source and sink nodes. This work was published in Smith et al. (2020).

Chapter 4 investigates some of the dynamical properties of the swing equation, a nonlinear dynamical model of power flow. This model will be used in chapter 5 to develop a more sophisticated model of cascading failures with which to analyse network resilience. Chapter 4 therefore serves as a prelude to chapter 5, and highlights the model's important characteristics.

Chapter 5 uses the dynamical swing-equation based model of cascades to analyse network resilience. As in previous chapters, the resilience will be computed as a function node behaviour. The results in this chapter will then be contrasted with those of chapter 3 which uses a steady-state cascade model.

Chapter 6 develops a model of small, domestic scale renewable power grids. Power consumption data from UK households, as well as real photo-voltaic power production data will be used to inform these models and provide a realistic notion of network node heterogeneity. The methods developed throughout the thesis to analyse the impact of node heterogeneity on network efficiency and resilience will be demonstrated in these real-world data driven models.

Finally, chapter 7 summarises the thesis and highlights directions for future research.

# Chapter 1

# Introduction and literature review

## 1.1 Introduction

Networks comprise individual elements and the connections between them. Many systems arising in nature, engineering, and the social sciences can be described as networks. For example, the internet (De Domenico and Arenas, 2017), social groups (Kitsak et al., 2010; Borge-Holthoefer and Moreno, 2012), neurons in biological tissue (e.g. Bullmore and Sporns, 2009), the interaction of proteins inside cells (e.g. Jeong et al., 2001), and electrical distribution grids (Witthaut and Timme, 2012). Although the individual elements comprising these networks may often be quite simple, their interactions can give rise to complex, emergent behaviour (Bar-Yam, 2019). Providing quantitative and qualitative descriptions of these complex systems provides a formidable set of challenges.

Two important notions in network science are those of efficiency and resilience. For example, efficiency might gauge how much power is lost in an electrical network. Resilience might measure a network's ability to function coherently after some initial shock. Efficiency and resilience are typically analysed as a function of a network's topology, in order to find connectivity patterns

that optimise the system's performance. So far in the literature, relatively little attention has been given to the role of node behaviour in determining efficiency and resilience. This thesis bridges that gap by analysing networks whose nodes have heterogeneous behaviour. Each node may be either a source or a sink of some physical flow that travels across the network. Nodes may also have some ability to adjust their behaviour in reaction to the state of the network. The primary motivation for this investigation is the increasing proliferation of renewable energy sources in electrical networks. This trend is causing an upheaval in the traditional structure of power grids, producing power networks with variable numbers of generators and consumers (Parhizi et al., 2015). The proportions of generators and consumers may change throughout a day, or even be controlled by a central planner. Uncovering the dependence of efficiency and resilience on node behaviour is therefore essential for understanding this modern energy landscape.

The main objective of this thesis is therefore to investigate the efficiency and resilience of flows through networks whose nodes have heterogeneous behaviour. Techniques from a variety of disciplines will be used in pursuit of this goal. Game theory will be used to analyse flow routing efficiency. Power flows will be modelled as nonlinear dynamical systems. Crucially, network science will provide the means to investigate the interaction between the flows and the networks underpinning them. The rest of this chapter reviews the relevant literature in these disciplines, focusing on specific elements of each that will be utilised throughout the thesis. Section 1.2 gives a brief overview of network science and highlights the network models that will be used repeatedly in subsequent chapters. Section 1.3 reviews the literature on the so-called Price of Anarchy, a game theoretic measurement of inefficiency that will be the focus of chapter 2. Section 1.4 reviews network resilience and cascading failures. Section 1.5 reviews network power flow models. Finally, section 1.6 will summarise the chapter and set out the objectives for the rest of the thesis.

## 1.2 Networks

This section provides a brief review of network science. The aim is not to provide an exhaustive survey of the subject's literature, which is vast, but to outline the main models and concepts that will be used in this thesis. For a broader overview of network science and its fundamentals, the reader might refer to the textbook by Barabási et al. (2016).

The objective of network science is to provide researchers with the tools to analyse the complexity of interconnected systems. The field draws on a foundation of nonlinear dynamics, statistical physics, and graph theory. It has grown into a vibrant and interdisciplinary subject over the past two decades, driven in part by technological advances giving researchers vast amounts of data on real world networks (Tan et al., 2013). The field is now employed in a diverse set of roles, supporting important efforts such as the development of new pharmaceutical therapies (Maron et al., 2020; Cheng et al., 2019), maintenance and design of communications infrastructure (Milanović and Zhu, 2018), and epidemic modelling (see for example Pastor-Satorras and Vespignani, 2001; Iannelli et al., 2017; Steinegger et al., 2020). Indeed, the Covid-19 crisis has seen network science attain an unprecedented level of importance and scrutiny, having formed the backbone of many modelling projects in response to the pandemic (Arenas et al., 2020; Vespignani et al., 2020; Aleta et al., 2020; Estrada, 2020). With the proliferation of contact tracing apps (Ferretti et al., 2020; Ahmed et al., 2020) and network visualisations in the media (BBC, 2020; The New York Times, 2020; The Guardian, 2020a;b; Metro, 2020), the language and imagery of network science is increasingly entering the public consciousness.

Mathematically, a network is a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ comprising of a set of $n$ nodes $\mathcal{V}$ and $m$ edges $\mathcal{E}$. Two nodes are adjacent if there is a single edge connecting them. A network is said to be connected if any node can be reached from any other. Additionally, a network is said to be simple if there is no more than one edge connecting any two nodes, and if no node is adjacent to itself via a self-loop. Non-simple graphs may be referred to as multi-graphs. A graph may also have other properties, such as edge weights describing the strength

of the connections, or an orientation describing the direction of edges. The degree $K$ of a node is the number of other nodes to which it is adjacent.

The topology, or structure, of a network can be encoded in the form of matrices. Let us now summarise the most important of these. The *adjacency matrix* $A$ of a simple undirected network is a symmetric $n \times n$ matrix defined by

$$A_{ij} = \begin{cases} 1, & \text{if node } i \text{ is adjacent to node } j, \\ 0, & \text{otherwise.} \end{cases} \tag{1.1}$$

The degree $K$ of node $i$ is given by the $i^{\text{th}}$ row sum of $A$. If the network is non-simple, then its adjacency matrix is

$$A_{ij}^{\text{multi}} = \begin{cases} l, & \text{if } l \text{ edges link nodes } i \text{ and } j, \\ 0, & \text{otherwise.} \end{cases} \tag{1.2}$$

The *incidence matrix* $E$ is an $n \times m$ matrix describing the relationship between nodes and edges:

$$E_{ij} = \begin{cases} 1, & \text{if edge } j \text{ is incident on node } i, \\ 0, & \text{otherwise.} \end{cases} \tag{1.3}$$

If the network is directed, then its incidence matrix will have signed entries:

$$E_{ij}^{\text{dir}} = \begin{cases} 1, & \text{if edge } j \text{ originates at node } i, \\ -1, & \text{if edge } j \text{ terminates as node } i, \\ 0, & \text{otherwise.} \end{cases} \tag{1.4}$$

Another important matrix is the *Laplacian* $\mathcal{L}$, the $n \times n$ matrix defined as

$$\mathcal{L}_{ij} = \begin{cases} \sum_{k=1}^{n} A_{ik}, & \text{if } i = j, \\ -A_{ij}, & \text{if } i \neq j. \end{cases} \tag{1.5}$$

The Laplacian can be computed as

$$\mathcal{L} = D - A, \tag{1.6}$$

where $D$ is the *degree matrix*, a diagonal matrix containing the degrees of each node. The Laplacian matrix can also be computed as

$$\mathcal{L} = EE^T. \tag{1.7}$$

If the network is weighted, then an equivalent weighted Laplacian is given by

$$\mathcal{L}_W = EWE^T, \tag{1.8}$$

where $W$ is an $m \times m$ diagonal matrix containing the weights of each edge. Many of the techniques of network science then derive from analysing the above matrices. Of particular importance is the eigenvalue spectrum of $\mathcal{L}$, which, amongst other things, gives information about the connectivity of the network. In particular, if the network is connected then $\mathcal{L}$ will have precisely one eigenvalue equal to 0. If a network has split into two distinct *connected-components*, which are internally connected but not connected to each other, then $\mathcal{L}$ will have two eigenvalues equal to 0. In general, the multiplicity of the 0 eigenvalue gives the number of connected-components. For a proof of this, see for instance Manik et al. (2014). Beyond giving a notion of connectivity, this eigenvalue property also introduces some numerical difficulties which will be touched upon at several points in this thesis.

Network scientists have developed an extensive array of techniques to characterise the structural properties of networks. In addition to measurements of network connectivity, an important notion is that of centrality. Many variants of and methods to quantify centrality have been devised, all of them providing a measure of the importance of particular nodes and edges in the structure of a network. One of the earliest was Katz centrality (Katz, 1953), which counts the number of possible walks from a given node to any other node. A sum is formed, weighted to give short walks a larger contribution. The higher the value of this sum, the more central the node. PageRank (Page et al., 1999) and eigenvalue-centrality (Bonacich, 2007) are closely related. There is also betweenness centrality, which computes all the shortest paths between all possible pairs of nodes and then computes the fraction of these paths that pass through a given node (Brandes, 2001). The length of paths within networks is also an important feature. This is often measured

using the characteristic path length (Watts and Strogatz, 1998), which is the mean length of all shortest paths between each pair of nodes in the network. Also related is communicability, which takes into account both shortest paths and longer deviations between pairs of nodes to interpret how easily information can be transmitted from one to the other (Estrada and Hatano, 2008). Another crucial network attribute is that of clustering. This is often quantified using the clustering-coefficient (Watts and Strogatz, 1998), which measures how connected a given node's neighbours are to each other.

Perhaps the most fundamental characteristic of a network is the degree distribution of its nodes. The nodes in a regular network, or lattice, will all have constant degree $K$. However, many networks in the real world do not have any such regular structure. The degrees of these more complex networks may be modelled as random variables. The properties of the resulting distributions can be used to categorise network structures. For example, a network whose nodes are connected together completely at random will have a binomial or Poisson degree distribution. Other networks may have a power-law distribution, indicating that the vast majority of nodes have low degree with a few nodes of much higher degree acting as centralised hubs. Various methods have been devised to artificially generate random networks with specific degree distributions. Not only do these methods furnish researchers with the ability to create synthetic networks resembling real-world structures, they also provide intuition about how the complexity of these structures can arise naturally from simple principles. Let us now review the three most common such methods, each producing a distinct type of network. They will be used throughout this thesis to provide a test-bed of synthetic networks from which general results and conclusions will be derived.

The simplest of these methods is the Erdős–Rényi procedure (Erdős and Rényi, 1960). Starting with $n$ nodes and no edges, the method then connects each pair of nodes with a probability $p$. The resulting networks have binomially distributed degrees. The case of $p = 1$ produces the complete graph with constant degree $K = n-1$. The Erdős–Rényi method can in fact be regarded as a percolation process on the complete graph (see Albert and Barabási, 2002, for more details). Most networks occurring in the real world are sparse

with mean degree $\overline{K} \ll n$. For these networks, the binomial distribution can be well approximated by the Poisson distribution. However, the Poisson distributed nature of Erdős–Rényi networks is typically unable to replicate the features of many networks found in nature. For instance, the degrees of the nodes in a Poisson network are relatively tightly grouped near the mean degree. This means the networks are unlikely to contain many large hubs or peripheral outliers. Poisson networks also possess very low clustering.

The second of the three main methods is the Watts–Strogatz algorithm (Watts and Strogatz, 1998), illustrated in figure 1.1. This method is able to

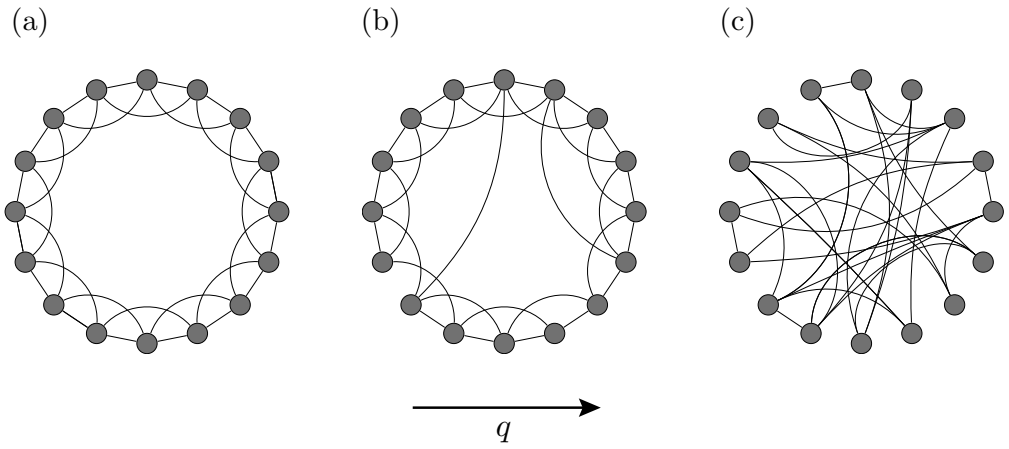



Figure 1.1: Illustration of the Watts–Strogatz algorithm (Watts and Strogatz, 1998). Starting with a ring lattice, such as the one shown in (a), the algorithm rewires the edges with a probability $q$. The network in (b) shows a realisation of a Watts–Strogatz graph with $q = 0.1$ and is typical of a small-world network. (c) shows an instance where $q = 1$, giving a Poisson distributed network.

capture some of the key properties of real networks, in particular the so-called small-world property. A small-world network is one in which the characteristic path length is small, like in a Poisson network, but the clustering coefficient is high like a regular lattice. This small-world quality can be observed in social networks for example (Barabási et al., 2016). The Watts–Strogatz method constructs small-world networks in the following way. First, start with a regular ring lattice, as shown in figure 1.1(a), with each node connected to its $K/2$ nearest neighbours. Each edge in the lattice is then selected and randomly rewired to another node with a probability $q$. As $q$ increases from 0,

the characteristic path length decreases sharply while the clustering coefficient remains approximately constant until $q \approx 0.1$ after which it decreases. The small-world networks occupy this region of approximately $q = 0$ to $q = 0.1$. An example is illustrated in figure 1.1(b), showing the long range shortcuts that typify small-world networks and result in the high cluster, low path length quality. At $q = 1$, the networks are completely random and we recover the Poisson case where both the clustering and path lengths are small. The Watts–Strogatz method will be used multiple times in this thesis, with the parameter $q$ being used to interpolate between regular lattices at $q = 0$ and Poisson networks at $q = 1$.

The final method we will consider is that of Barabási and Albert (1999) which is used to construct so-called scale-free networks. These networks possess power-law degree distributions. This means that most nodes will have a low degree, but there will be a few highly connected hubs. This hub and spoke structure, and the associated power-law degree distribution, can be observed in many settings such as airline route networks and the internet (Albert and Barabási, 2002). The Barabási–Albert method constructs scale-free networks using preferential attachment. The method starts with a small cluster of only a few nodes. A new node is then added and connected to the pre-existing nodes with a probability proportional to their degrees. New nodes are added in this fashion until the desired network size $n$ is reached. Preferential attachment is also referred as the "rich-get-richer" principle and encapsulates the idea that highly connected nodes have an enhanced ability to create yet more connections, leading to centralised hubs. This can be seen in social networks where popular socialites can continue to accrete friends via their pre-existing connections, while the less popular members of the network remain isolated.

Figure 1.2 summarises these key network types — Poisson, small-world, and scale-free — and the methods used to create them. There are many other types of network and various algorithms to create them. For example, multiplex and multilayer networks. These represent systems with multiple different but coupled types of interactions. Refer to De Domenico et al. (2013) for a mathematical introduction to multiplex networks. However the methods

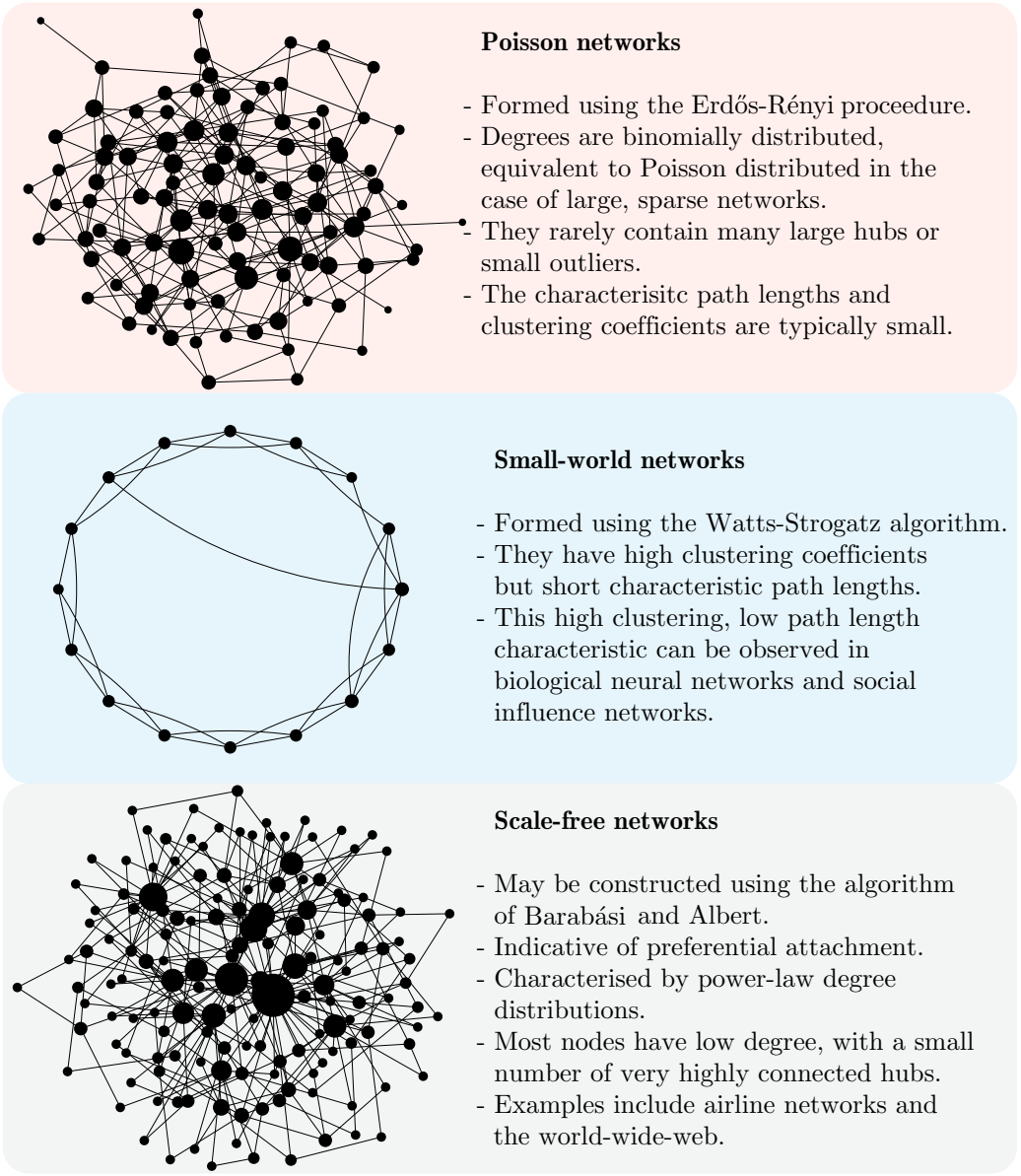



Figure 1.2: Summary of three important network types. The sizes of the nodes in the illustrations indicate their degrees. The network visualisations were made using the Networkx (Hagberg et al., 2008) package in Python.

of figure 1.2 will form the basis for constructing synthetic networks in much of the following work.

Networks may also have some dynamics. The dynamics may concern the topology itself, with the structure of the network evolving through time (e.g. Bianconi and Barabási, 2001; Bornholdt and Rohlf, 2000). The dynamics

could alternatively concern some property or quantity of the nodes and edges, for example some physical quantity such as traffic flowing through the network (e.g. Coclite et al., 2005). Sometimes these two types of dynamics may even be coupled, with the flow through a network influencing the network's structure. For an example of this see Garlaschelli et al. (2007). An extensively studied form of network dynamics that of coupled oscillators, the most famous model of which is that of Kuramoto (1975). For a survey of oscillator dynamics on networks refer to Dörfler and Bullo (2014). The use of coupled oscillators is widespread within mathematical biology, where they are employed to study networks of neurons. Various different coupled ODE models are used in the literature to study biological networks of neurons, such as non-smooth and piece-wise linear models (e.g. Nicks et al., 2018; Coombes et al., 2018), the stability of which can be interrogated using the master stability function (Coombes and Thul, 2016). Recently, oscillator dynamics have been used to model electrical networks (Witthaut and Timme, 2012), a topic which will be explored further in section 1.5.2. Other dynamical processes on networks include Ising models (Hopfield, 1982), percolation (e.g. Cellai et al., 2013), and epidemic spreading (e.g. Pastor-Satorras et al., 2015). Discrete versions of differential operators can also be posed on networks, such as the diffusion equation (e.g. Gomez et al., 2013; Angstmann et al., 2013). The general form of diffusion on a network is

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -\mathcal{L}x(t) + s(t), \tag{1.9}$$

where $x \in \mathbb{R}^n$ is a vector of some quantity associated with each node, $\mathcal{L}$ is the Laplacian and $s \in \mathbb{R}^n$ is a source term. Variants of equation (1.9), such as pattern forming reaction-diffusion systems, have also been studied on networks (Kouvaris et al., 2015). Another important category of network dynamics is that of game theory. For example, Gómez-Gardeñes et al. (2012) looked at the prisoners dilemma being played by nodes on a multiplex network with best response dynamics.

This thesis is primarily concerned with flows through networks, and their efficiency and resilience. Network dynamics are central to this investigation. The network flows themselves will be modelled using various dynamical

methods. For instance, chapter 2 draws upon game theory to compute flows and assess their efficiency, while chapter 3 models flows that are equivalent to steady states of equation (1.9). Chapters 4, 5 and 6 analyse electrical power flows using a type of nonlinear oscillator model. Additionally, network cascades, a form of topological dynamics coupled to flow dynamics, will be used in chapters 3 onwards to assess resilience. Each of these aspects will be explored in detail in the remainder of this literature review.

## 1.3 The Price of Anarchy $\mathcal{P}$

In chapter 2, the efficiency and redundancy of physical flows in networks will be quantified using a measurement from game theory called the Price of Anarchy $\mathcal{P}$. Introduced by Koutsoupias and Papadimitriou (1999), $\mathcal{P}$ gauges the inefficiency caused by a lack of cooperation in competitive systems. This section introduces $\mathcal{P}$ and surveys the attendant literature. We begin by reviewing some key concepts from game theory that are necessary to provide a definition of $\mathcal{P}$.

### 1.3.1 Game theory and definition of $\mathcal{P}$

The modern discipline of game theory began with the foundational work of Morgenstern and Von Neumann (1944). It provides a mathematical framework in which to analyse competitive systems. Typically, a game in this sense will comprise $n$ players competing with each other to maximise their payoffs. Examples include animals competing for food, and businesses competing to maximise profit. A player is said to be rational if they always act to increase their payoff, if such a move is possible. If the players are rational and do not cooperate with each other, then the system will eventually reach a state called a Nash equilibrium. These equilibria are defined as states where no player can increase their payoff by altering their strategy, given the strategies of the other players. Each non-cooperative game may have many Nash equilibria. Since being introduced by John Nash (1950; 1951), the

analysis and computation of Nash equilibria has become fundamental to game theory. They are in general sub-optimal; if the players are instead able to cooperate and coordinate their strategies, then they can often achieve higher overall payoffs. The optimal state that can be reached through cooperation will be referred to as the system optimum (SO). The SO maximises the total payoff to all players, and by definition it will yield a total payoff greater than or equal to that of any Nash equilibrium. A well studied example in which non-cooperative behaviour leads to a sub-optimal Nash equilibrium is that of road traffic, where drivers are effectively in competition for routes of shortest travel time. Drivers do not typically cooperate with other drivers when choosing their routes. This is often referred to as "selfish routing" and leads to congestion that might be avoidable if the drivers instead coordinated their efforts (Roughgarden and Tardós, 2002). Another example is the "tragedy of the commons" (Lloyd, 1833; Hardin, 1968), in which villagers vie to graze their livestock on an area of common land. If the villagers act non-cooperatively, they will all use as much of the commons as possible at any given time. It can be seen that this situation is a Nash equilibrium by considering that the only change of strategy available to any villager is to use less of the commons. If they do this, then the other villagers will simply move to occupy the newly available land since this increases their payoffs. Therefore any change of strategy simply decreases the villager's payoff, given the strategies of the other players. This Nash equilibrium behaviour eventually leads to the destruction of the commons; a situation that could have been avoided had the villagers coordinated to ration the land. In economics, the tragedy of the commons has come to refer generally to the depletion of a common resource through non-cooperative behaviour. Examples include the overfishing of marine stocks (McWhinnie, 2009), the over-ploughing of soil resulting in the dust bowl of 1930's Oklahoma (Lal, 2009), and the panic buying of toilet roll (Loxton et al., 2020).

We have so far considered games as systems in which players seek to maximise a payoff. A reciprocal type of game can be formulated in which players instead compete to minimise a cost. This is the usual fomulation of network traffic routing games, which will form the foundation of chapter 2. Let us now work towards a formal definition of the Nash equilibrium and the system optimum

in the context of a cost-minimisation game. Let the game consist of $n$ rational, non-cooperative players. Each player will have a set of possible strategies. Let $S_i$ denote the $i^{\text{th}}$ player's set of possible strategies, with $s_i \in S_i$ being their particular choice of strategy. The strategy profile of the game is the $n$-tuple containing the strategy choices of all players:

$$s = (s_1, s_2, \ldots, s_n) \quad \in \quad S = S_1 \times S_2 \times \cdots \times S_n. \tag{1.10}$$

Each player $i$ has a cost function $c_i(s)$, which computes the cost they will incur if the strategy profile $s$ is played. Note that the cost to player $i$ is a function of all the players' strategies, not just their own. Now let $s_{-i}$ denote the strategy profile for all players except for player $i$:

$$s_{-i} = (s_1, s_2, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n) \ \in \ S_{-i}, \tag{1.11}$$

where

$$S_{-i} = S_1 \times S_2 \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots \times S_n. \tag{1.12}$$

Finally, let $c_i(s_i, s_{-i})$ denote the cost to player $i$ if they adopt strategy $s_i$, given that all other players adopt strategy profile $s_{-i}$. We can now define a Nash equilibrium $s^*$ as a strategy profile

$$s^* = (s_1^*, s_2^*, \ldots, s_n^*) \ \in S, \tag{1.13}$$

such that for every player $i$

$$c_i(s_i^*, s_{-i}^*) \leq c_i(s_i, s_{-i}^*) \quad \forall s_i \in S_i. \tag{1.14}$$

This means that at $s^*$ no player can decrease their cost, given the strategies of the others. For a proof of the guaranteed existence of these equilibria in non-cooperative games, refer to Nash (1951). For a broader overview of Nash equilibria, including variants such as mixed-strategy Nash equilibria, the reader might refer to the book by Tadelis (2013). We can now also formally define the system optimum. It is the strategy profile $s^{\text{SO}} \in S$ that minimises the total cost:

$$\min_{s \in S} \sum_{i=1}^{n} c_i(s). \tag{1.15}$$

$s^{\mathrm{SO}}$ does not necessarily satisfy condition (1.14). This means that some players at the SO may be able to change strategies to decrease their cost. This subset of players must resist such changes, and thus tolerate an elevated cost, in order to minimise the total cost across the system as a whole. This sort of self-sacrifice for the greater good might be possible with cooperation or system-wide coordination. However, as we will see in chapter 2, some systems cannot behave in a cooperative way. These systems can sometimes be modified in order to drive their Nash equilibria towards optimality.

The Price of Anarchy $\mathcal{P}$ quantifies the sub-optimality of Nash equilibria. $\mathcal{P}$ is defined as the ratio of the total cost of a game's worst possible Nash equilibrium, to the total cost of its SO:

$$\mathcal{P} := \frac{\sum_{i=1}^{n} c_i(s^*)}{\sum_{i=1}^{n} c_i(s^{\mathrm{SO}})}. \tag{1.16}$$

$\mathcal{P}$ can be interpreted as measuring the cost of non-cooperation, since the SO is typically reached through cooperation, and the Nash equilibrium through a lack of it. The higher the value of $\mathcal{P}$, the greater the inefficiency caused by the players' non-cooperation. $\mathcal{P}$ is sometimes referred to in the literature as gauging the price of "selfishness" (Roughgarden, 2005). The players in a non-cooperative game can also be thought of as optimising only locally, whereas the SO is reached through a coordinated system-wide optimisation. Another interpretation of $\mathcal{P}$ is therefore that it measures the discrepancy between local and global optimisation in a competitive system.

### 1.3.2 Examples of $\mathcal{P}$ in competitive systems

Chapter 3 uses $\mathcal{P}$ to gauge the efficiency of least resistance flows in networks, by exploiting their equivalence with the Nash equilibira of routing games. This work is mostly derived from previous studies of $\mathcal{P}$ in relation to network traffic flows, but let us first review of some of the other contexts in which $\mathcal{P}$ has been investigated. One of the most common applications is in network creation games, where the nodes of a network seek to connect up to each other by sprouting edges. A popular variety of this game was introduced by

Fabrikant et al. (2003). In this game, every node pays a cost $c$ for each edge it creates. The objective of each node is to be connected to all other nodes whilst minimising the sum of the costs. Fabrikant et al. showed that all Nash equilibria in these games result in tree networks, and that any network that is not a tree is just a transient state in the game. They also argued that if a network is a tree then $\mathcal{P}$ is constant with respect to $c$. These results were later revised by Albers et al. (2006) who showed that for certain values of $c$, there can in fact be non-transient Nash equilibria giving networks containing cycles. They also demonstrated that $\mathcal{P}$ decreases with $c$, as well as providing upper bounds on $\mathcal{P}$. Upper bounds on $\mathcal{P}$ for other variants of the network creation game have also been derived (Demaine et al., 2007).

Another important application of $\mathcal{P}$ is in machine load balancing games, such as the allocation of resources in batch computing processes. The usual formulation of these games is that a job $i$ needs to choose a machine $j$, adding a load $w_j(i)$ to that machine. The jobs seek to match to machines with the lowest possible load. Various upper bounds for this class of game have been proved, such as $\mathcal{P}$ being bounded by $\mathcal{O}(\log m/\log\log m)$ for $m$ machines (Fiat et al., 2007).

Knight and Harper (2013) used $\mathcal{P}$ to measure the inefficiency caused when NHS patients are free to choose the location of their treatment. They showed that allowing this choice leads to a marked decrease in efficiency, particularly in health care services that ordinarily have sufficient capacity. Similarly, Rose (2018) studied $\mathcal{P}$ in a customer-to-service matching game, and applied the analysis to the case of Uber drivers matching with passengers.

### 1.3.3 $\mathcal{P}$ in network flow routing games

Much of the work on $\mathcal{P}$ in flow routing games can be traced back to work of Roughgarden and Tardós and their work on "selfish routing" (Roughgarden and Tardós, 2002; Roughgarden, 2005). In these papers, they examine traffic flowing through a network from a source node to a sink node. Each unit of traffic incurs a cost representing travel time for each edge they traverse. The objective of each unit is to minimise their total travel time from source to

sink. Roughgarden and Tardós (2002) approximate the traffic flow on each edge as a continuous quantity, since each unit of traffic makes up only a very small amount of the overall traffic volume. This so-called non-atomic flow approximation has been used widely in traffic flow modelling (Coclite et al., 2005) and in subsequent publications on the Price of Anarchy (Roughgarden, 2003; 2006; Skinner, 2015; Rose et al., 2016; Youn et al., 2008). Each edge $e$ in the network then possesses a continuous flow volume $f_e \in \mathbb{R}$, and a cost function $c_e(f_e)$ giving the travel time. These cost functions are typically monotonically increasing functions of $f_e$. This means the edges are congestible; the more traffic that uses them, the longer the travel time becomes. Each infinitesimal unit of traffic competes non-cooperatively with the rest of the traffic for paths of least cost. The resulting Nash equilibrium is achieved when all routes between the source and the sink have equal cost, meaning that none of the traffic can change its routing strategy to decrease its cost. The total cost across the network is

$$\mathcal{C}(f) = \sum_{i=1}^{m} c_i(f_i) f_i, \tag{1.17}$$

where $m$ is the number of edges and $f \in \mathbb{R}^m$ is the flow vector giving the flow volumes on each edge. The system optimum is the traffic pattern $f^{\mathrm{SO}}$ that minimises this total cost. $\mathcal{P}$ is then given by

$$\mathcal{P} = \frac{\mathcal{C}(f^{\mathrm{Nash}})}{\mathcal{C}(f^{\mathrm{SO}})}. \tag{1.18}$$

The fundamental example given by Roughgarden and Tardós (2002) is the network depicted in figure 1.3(a), which was first studied by Pigou (1920). A flow volume of 10 enters the network at the source node on the left and travels to the sink node on the right. The nodes are connected by two edges. The bottom edge has a fixed cost of 10, indicated in figure 1.3(a) by the cost function $c_2(f_2) = 10$. The top edge has a variable cost which is equal to its traffic volume: $c_1(f_1) = f_1$. This top edge can be thought of as a congestible short-cut. If little traffic takes this edge, it will have a low cost and provide a cheap route from source to sink. At the Nash equilibrium, illustrated in figure 1.3(b), all 10 units of traffic volume take this top edge so that its cost is $c_1(10) = 10$. Both routes therefore have an equal cost with

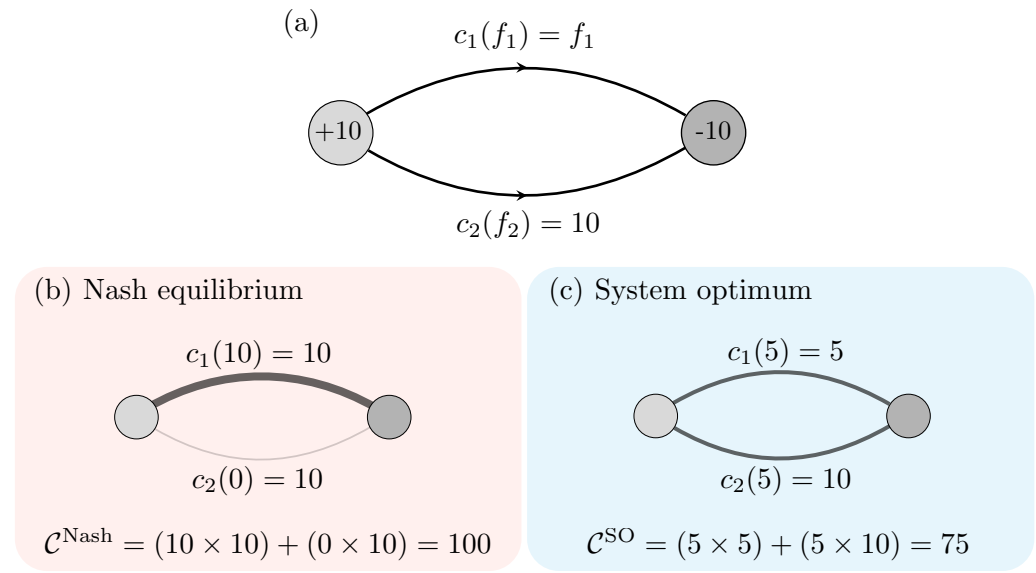



Figure 1.3: (a) The Pigou network. 10 units of traffic flow volume enter at the source node on the left and travel to the sink node on the right. The top edge has variable cost and the bottom edge has a fixed cost of 10. (b) At the Nash equilibirum, all of the flow take the top edge giving a total cost of 100. Edge thickness indicates relative traffic usage. (c) At the SO, half of the flow takes the top edge and half takes the bottom giving a total cost of 75. The resulting Price of Anarchy is $\mathcal{P} = 100/75 = 4/3$.

$c_1 = c_2 = 10$. This is a tragedy of the commons situation; so much of the traffic has opted to take the potentially cheaper route that its cost is now as high as the fixed cost route. The total cost in the network, computed via equation (1.17), is then $\mathcal{C}^{\text{Nash}} = 100$. It is straightforward to show that the total cost is minimised, giving the SO, when only half the flow takes the variable cost edge. This is illustrated in figure 1.3(c) and gives a total cost of $\mathcal{C}^{\text{SO}} = 75$. Consequently, $\mathcal{P} = \mathcal{C}^{\text{Nash}}/\mathcal{C}^{\text{SO}} = 4/3$. Roughgarden and Tardós (2002) showed that 4/3 is in fact the upper bound on $\mathcal{P}$ for any network with linear cost functions. They also provided upper bounds for networks with quadratic and cubic cost functions. In a subsequent paper (Roughgarden, 2003) it was shown that these upper bounds are independent of network topology, and depend only upon the type of edge cost functions.

Roughgarden (2006) showed that there is no efficient algorithm to find the optimum network for a non-cooperative routing problem. Given this algorithmic complexity, it may be more fruitful to identify the general topological

characteristics that tend to produce inefficient Nash equilibria. This might then inform the design of road networks. In this vein, Youn et al. (2008) simulated non-cooperative traffic flowing through the road networks of Boston, Manhattan and London. For edge cost prices they used the Public Bureau of Roads function (Bureau of Public Roads, 1964)

$$c_e(f_e) = \frac{d_e}{v_e}\left[1 + \gamma\left(\frac{f_e}{p_e}\right)^{\mu}\right], \tag{1.19}$$

where $d_e$, $v_e$ and $p_e$ are the length, speed limit, and number of lanes respectively of road $e$. $\gamma$ and $\mu$ are some empirical constants derived from road traffic data. In each network, they found that $\mathcal{P}$ exhibits a series of ripples as the total traffic flow volume is increased, and that $\mathcal{P}$ never reaches the upper bound. The morphology of these ripples were found to be unique to each network topology. As an abstract model, they then simulated traffic through large ensembles of small world, Erdős-Rényi and scale-free networks with linear edge cost functions given by

$$c_e(f_e) = \alpha_e f_e + \beta_e. \tag{1.20}$$

The coefficients $(\alpha_e, \beta_e) \in [1, 3] \times [1, 100]$ were chosen at random for each edge. They calculated the ensemble average $\mathcal{P}$ as a function of the total flow volume for these networks, and found that in all cases $\mathcal{P}$ smoothly increases to a maximum whose location is a function of the mean values of the coefficients in equation (1.20). In the high volume limit, $\mathcal{P}$ decreases to 1. This is due to the networks becoming so congested that both the Nash and SO strategies are equally costly.

Skinner (2015) studied the more abstract model of a square lattice network, with flow entering on one side and being routed to the other. Some edges in the lattice are chosen at random to have a variable cost, and the rest to have a fixed cost. In particular, the cost functions of the edges are chosen to be $c_e(f_e) = f_e$ with a probability $p$ and $c_e(f_e) = 1$ with a probability $1 - p$. At $p = 0$, the lattice will consist entirely of fixed cost edges, and at $p = 1$ entirely of variable cost edges. Skinner showed that $\mathcal{P}$ can be maximised when $p$ is equal to the percolation threshold of the lattice. Rose et al. (2016)

built on this work and showed that $\mathcal{P}$ also depends on the aspect ratio of the lattice and the total flow volume $F$, and that the maximum does not always coincide with the percolation threshold. It was shown that $\mathcal{P}$ exhibits a series of ripples as a function of $p$ and $F$, similar to those found by Youn et al. (2008). Rose et al. showed that the locations of these ripples with respect to $p$ and $F$ correspond to critical points at which it becomes beneficial for the flow to travel down more expensive routes, as the previous best routes become congested through overuse.

Aside from manipulating network topology, composition, and flow volume, it might be possible to increase efficiency by changing the behaviour of the flow. Balcan et al. (2013) demonstrated that since the non-cooperative flow tends only to act on local knowledge, it can be advantageous to feed the traffic flow some global information. For example, a central planner or traffic authority could provide information to drivers about the state of other parts of the network. They show that even small amounts of global knowledge will cause the flow to quickly approach the best possible Nash equilibrium. Çolak et al. (2016) looked at a traffic routing model whose constituents have a "social good" parameter, allowing the behaviour of the flow to be smoothly interpolated from completely selfish to completely cooperative. It was found that only a small increase in cooperative behaviour leads to substantial gains in efficiency. Encouragingly, these two papers show that by modifying the flow behaviour only slightly, one can obtain a much quicker and more substantial increase in efficiency compared to the difficult methods of manipulating the network topology itself. Such modification of traffic is becoming increasable feasible due to smart traffic routing apps, variable speed limits and variable road pricing. However, Cole et al. (2006) showed that finding optimal pricing strategies is NP-hard, and no variable pricing strategy can offer any benefits greater than those available by removing or rewiring edges. A further drawback to flow behaviour modification is that drivers may be inherently irrational. Lima et al. (2016) found that traffic flow patterns often deviate far away from the Nash flow pattern, due to factors like drivers having favourite routes. Of course it may not be possible in many cases to have any influence over the fundamental behaviour of the flow, as is the case for electrical flows. In this situation we are confined to change

network properties alone.

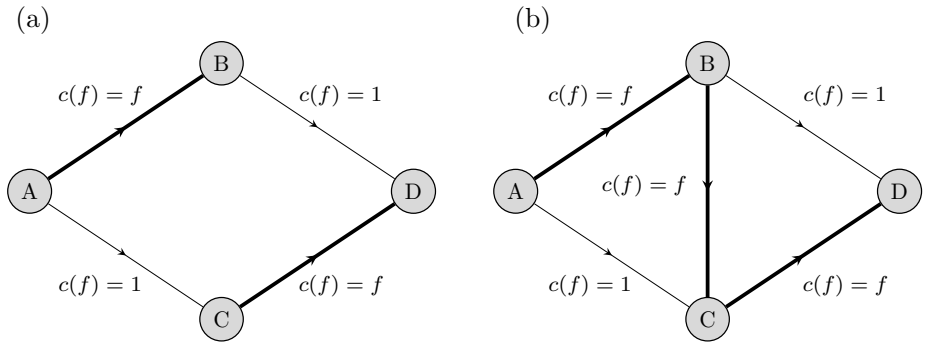



Figure 1.4: Demonstration of Braess's paradox. In both (a) and (b), flow enters the network at node A, and exits at node D. Both networks comprise a mixture of congestible variable cost edges, with $c_e(f_e) = f_e$, and non-congestible fixed cost edges with $c_e(f_e) = 1$. In (a) the two available routes are ABD and ACD, both of which have the same composition of variable and fixed cost edges. This means there is no advantage to choosing one route over the other. The Nash flow is therefore identical to the SO, giving $\mathcal{P} = 1$. In (b), a new variable cost edge BC is introduced. This gives an additional route ABCD, which consists only of variable edges. As much selfish flow as possible attempts to use this new route, leading to increased congestion and $\mathcal{P} > 1$. Network (b) is sometimes referred to as a Wheatstone bridge in electrical engineering.

In some cases the Nash flow may overuse a particular edge to such an extent that it would be better to simply remove it from the network. Such a system is illustrated in figure 1.4(b), where the removal of the central edge causes a decrease in the cost of the Nash equilibrium, and correspondingly a decrease in $\mathcal{P}$. This counter-intuitive behaviour is referred to as Braess's paradox (Braess et al., 2005), and can be broadly defined as a situation in which the addition of an edge to a network causes increased inefficiency due to over-competition for the new resource.

Braess's paradox has been studied in a wide range of settings. For example, in their work on $\mathcal{P}$ in road networks, Youn et al. (2008) found that removing certain roads decreased traffic times. Under their analysis, the removal of Blackfriars bridge in London would ease congestion. Cohen and Horowitz (1991) reported Braess's paradox in electrical networks where the cost on each

edge is a voltage drop. They found that the addition of an edge can cause an increase in power loss. Wang and Baillieul (2016) outlined a measurement called the "cost of link addition" to quantify the extent of the paradox in electrical networks. The measure is defined as the ratio of the power loss after a link is added, to the power loss before. That Braess's paradox can occur in electrical networks is perhaps unsurprising, given that the current is comprised of infinitesimal particles seeking paths of least resistance. The correspondence between electrical flows and the Nash equilibria of non-atomic routing games will be addressed in chapter 2.

The network in figure 1.4(b) is referred to as a Wheatsone bridge in electrical engineering. Milchtaich (2003) proved that Braess's paradox can only occur if there is at least one Wheatstone bridge contained within the network. Braess's paradox is often mentioned in connection with the Price of Anarchy. The two concepts are indeed related, however, $\mathcal{P}$ is not dependent upon the existence of Wheatstone bridges. For example, the Pigou network in figure 1.3 contains no Wheatstone bridges and yet has $\mathcal{P} = 4/3$. The correspondence between $\mathcal{P}$ and the inefficiency caused by excess capacity in a network will be looked at in detail in chapter 2.

More exotic variants of the paradox have also been reported. Solé-Ribalta et al. (2016) investigated transport on multiplex networks, where each layer is a different mode of transport. They show that in some circumstances the addition of more layers to the multiplex can increase congestion. Witthaut and Timme (2012) investigate Braess's paradox in networks of coupled oscillators. They show that the removal of an edge from the network can sometimes cause the critical coupling capacity to decrease, which means that the removal of the link has allowed the network to synchronise more easily. They explain their results by noting that the presence of small cycles within the networks can cause an effect similar to so-called geometric frustration seen in spin glasses. If removing a link causes the break up of a small cycle, then this may lead to a drop in the critical coupling capacity. In further work on oscillator network models, it was shown that introducing an additional link can cause a synchronous state to become unstable (Coletta and Jacquod, 2016). Oscillator models, particularly of electrical grids, will be reviewed

in further detail in this literature review in section 1.5.2. Oscillator models will then form the basis of chapters 4 to 6, where they will be used to model AC power flow and investigate the resilience of electrical networks to edge removals.

## 1.4 Network resilience

While chapter 2 investigates the efficiency of flow networks using the Price of Anarchy, subsequent chapters turn their attention to network resilience. This section provides a brief overview of the network resilience literature.

The resilience of a network is broadly defined as its ability to function after the removal of some of its structure. A common measure of resilience is to remove nodes or edges at random and record the fraction that must be removed to disconnect the network. The higher this fraction, the more resilient the network. This type of process can be formalised using the language of percolation theory. The book by Stauffer and Aharony (2018) provides a good general overview of percolation theory. A percolation process is where a node (or edge) is chosen to be "occupied" with a probability $p$. Occupied in this sense may mean that the node is activated or allowed to function properly. The percolation threshold $p_c$ is the point at which a giant connected component emerges that comprises only occupied edges. The inverse percolation process is where nodes (or edges) are chosen to be "de-occupied" or removed from the network with a probability $l = 1 - p$. The critical threshold $l_c$ is then the fraction of edges that need to be removed to disconnect the giant connected component of a network. $l_c$ is related to the regular percolation threshold by $l_c = 1 - p_c$. The higher the value of $l_c$ , the greater the number of nodes (or edges) that must be removed to disconnect the network and thus the more resilient is the network.

Percolation processes on networks have received a great deal of attention in the literature, and the reader might refer again to Barabási et al. (2016) for a high-level overview. Let us summarise some of the important results. The critical threshold on Erdős-Rényi networks can be computed exactly as

$l_c = 1 - 1/\overline{K}$. This means that the greater the mean degree of a Poissonian random network, the higher the value of $l_c$ and so higher the resilience. For example if a Poisson network has $\overline{K} = 3$, then 2/3 of its nodes need to be removed to completely disconnect the network. Percolation on small-world networks (Moore and Newman, 2000a;b) and scale-free networks (Cohen et al., 2002; Schwartz et al., 2002), have revealed them to possess very high values of $l_c$ making them robust against random removals. This resilience is due to the hub structure of the networks. A randomly removed node is most likely to be one of the many peripheral nodes, and not a highly central hub. However, these types of networks are extremely vulnerable to a targeted attack that deletes nodes with highest degree or centrality. Such a procedure can rapidly dismantle scale-free networks (Albert et al., 2000; Cohen et al., 2001).

Measuring the resilience of a network using this percolation methodology allows researchers to find ways of optimising network structure to boost resilience. For example Halu et al. (2016) used a simulating-annealing method to decrease the critical threshold $l_c$ of electrical grids. Importantly Paul et al. (2004) and Tanizawa et al. (2005) showed that networks with a bimodal degree distribution can be made resilient against both random and targeted attacks.

The percolation method only takes into account network topology. However, many networks also possess some form of dynamics. In this case, a measure of resilience should also gauge the network's ability to support the proper function of the dynamical process. For example, the removal of an airport from an airline network may not be sufficient to disconnect the network, but it may cause such congestion that the transportation network can no longer serve its purpose. This is an example of a cascading failure. Other examples include electrical blackouts (Carreras et al., 2002; Dobson et al., 2007), overloads in the Internet (Crucitti et al., 2004), and crashes in financial markets (Huang et al., 2013; Haldane and May, 2011). In these scenarios, a purely topological measure is not enough to capture resilience. In response to this problem, various ways of modelling cascading failures have been devised. The simplest is perhaps the sandpile model (Goh et al., 2003), where some

load builds up on a node until a maximum carrying threshold is exceeded. At this point, the node is taken out of operation and its load passed onto adjacent nodes, causing an avalanche of failures. This type of model is able to recreate the power-law distribution of cascade sizes typically observed in real world systems (Dobson et al., 2007). Another type of threshold method was employed by Watts (2002), where a node fails if a fraction $\varphi$ of its $K$ neighbours have also failed. A critical region of the $(\varphi, K)$ phase space was identified within which power-law distributed cascades occur. It was also shown that networks with more heterogeneous degree distributions are typically more resilient to cascades. A flow based cascade model was investigated by Motter and Lai (2003), Lai et al. (2004), and Crucitti et al. (2004) where a network flow is computed using a steady state model of electricity. An edge in the network was chosen to fail if it carries an amount of flow greater than the edge's threshold $\alpha$. It was shown that networks are particularly vulnerable if their edges carry a flow load proportional to the degree of the adjoining nodes. In this case, the cascade behaves like a targeted attack, taking out high degree hubs. A variant of this model will be developed in chapter 3 to probe the resilience of electrical grids with high numbers of small generators. Cascades have been studied on large scale power grid topologies in Pahwa et al. (2014), and in Yang et al. (2017) which showed that large scale cascades are most likely to be triggered by edges in the vicinity of a network's core. Cascades have also been studied in interdependent, multiplex networks (e.g. Zhang et al., 2018; Buldyrev et al., 2010). Various strategies to protect networks from cascades have been investigated. For example, selective pruning of a network topology was shown to help arrest cascades (Motter, 2004; Witthaut and Timme, 2015), whilst a rewiring scheme to suppress cascades was considered by Brummitt et al. (2012). Fang et al. (2015) considered a combinatorial optimisation scheme to enhance network topology for robustness against cascades.

Chapters 3 onwards will develop methods for measuring the impact of cascades on flow networks, particularly in relation to electrical grids. The impact of variable node function and increased power distribution of sources will be considered. This will highlight the possibility of altering a network's resilience without changing its connectivity.

## 1.5 Modelling electrical networks

Chapters 3 onwards turn their attention specifically to the resilience of electrical flow networks. In particular the modern trend of electrical networks that possess variable source and sink numbers. In electrical engineering there are various ways to compute electrical flow, with various advantages and disadvantages. Two methods are used in this thesis. The first is the steady state linearsed DC power flow method. The second is the so-called swing equation. This final section of the literature review looks at both of these equations and reviews their usage in the wider literature. This section assumes that the reader has some basic familiarity with the fundamentals of AC power flow, and terms such as impedance and susceptance. For a primer on AC power flow, one might refer to the textbook by Glover et al. (2016).

### 1.5.1 The DC power flow equation

Electrical circuits are governed by Kirchoff's laws, which specify conservation of current flow and voltage. For an overview of Kirchoff's laws, refer to the textbook on electrical circuits by Alexander and Sadiku (2012). The DC power flow equation is a steady state model for computing power flow in a network, and gives a direct solution to Kirchoff's laws. It is in essence a DC approximation of an AC power network. Note that in an AC setting, the total or apparent power $S$ is a complex number whose real and imaginary components are termed the real power $P$ and reactive power $Q$ respectively:

$$S = P + iQ. \tag{1.21}$$

The real power $P$ is the quantity that will be consumed by loads or consumers in an electrical grid, and is measured in Watts. It is this quantity that is important for our purposes and that we seek to model using the DC power flow equation. To derive the DC model for an electrical network with $n$ nodes

and $m$ edges, we start with the power flow equations:

$$P_i = |V_i| \sum_{j=1}^{n} |V_j| \left[ G_{ij} \cos\left(\theta_i - \theta_j\right) + B_{ij} \sin\left(\theta_i - \theta_j\right) \right], \tag{1.22}$$

$$Q_i = |V_i| \sum_{j=1}^{n} |V_j| \left[ G_{ij} \sin\left(\theta_i - \theta_j\right) + B_{ij} \cos\left(\theta_i - \theta_j\right) \right]. \tag{1.23}$$

$P_i$ and $Q_i$ are the real and reactive power respectively at node $i$, and $V_i$ is the voltage. The voltage phase angle at node $i$ is $\theta_i$. $G_{ij}$ and $B_{ij}$ are the conductance and susceptance respectively of edge $ij$. These are the real and imaginary parts of the admittance

$$Y_{ij} = G_{ij} + iB_{ij}, \tag{1.24}$$

where $i$ denotes the imaginary unit. The impedance of edge $ij$ is

$$Z_{ij} = R_{ij} + iX_{ij}, \tag{1.25}$$

where $R_{ij}$ and $X_{ij}$ are the resistance and reactance respectively. The impedance is the reciprocal of admittance:

$$Y_{ij} = \frac{1}{Z_{ij}}. \tag{1.26}$$

An in-depth discussion of the power flow equations is given in Li (2014). In power systems engineering, equations (1.22) and (1.23) are often solved iteratively via a Newton–Raphson method. However, under three reasonable assumptions for a power grid, the equations can be simplified to an $n$-dimensional linear system. The first assumption is that the reactance of the edges is much larger than the resistances: $X_{ij} \gg R_{ij}$. This allows the approximation $B_{ij} \approx 1/X_{ij}$. The second assumption is that the voltage phase differences $\theta_i - \theta_j$ are small, allowing small angle approximations in equation (1.22). The final assumption is that the magnitude of the node voltages across the network are roughly constant. Taking these three assumptions together, and normalising the voltages by the network's reference voltage

allows equation (1.22) to be approximated as

$$P_i = \sum_{j=1}^{n} B_{ij}(\theta_i - \theta_j). \tag{1.27}$$

Writing this in matrix-vector form gives the linearised DC power flow equation:

$$P = \mathcal{L}\,\theta, \tag{1.28}$$

where $P \in \mathbb{R}^n$ and $\theta \in \mathbb{R}^n$ are vectors containing the real power and voltage phase angle respectively at each node. $\mathcal{L}$ is the weighted graph Laplacian, also known as Kirchoff's matrix:

$$\mathcal{L} = E^T \mathrm{diag}(B)\, E, \tag{1.29}$$

where $B \in \mathbb{R}^m$ is the vector of susceptances.

The DC power flow equation (1.28) will be used in chapter 3 as a computationally efficient model of steady state power flow. This will then be used to investigate the resilience of electrical grids. A solution of equation (1.28) is in fact a steady state of the graph diffusion equation (1.9), and represents a solution to Kirchoff's laws. The equation is commonly used in the electrical engineering literature (Stott et al., 2009) and in popular commercial software such as ETAP. It is also employed more generally in network science. It was used by Motter and Lai (2003) to investigate the resilience of the US power grid. Halu et al. (2016) used it to model small scale power networks and optimise them for resilience. It also forms the basis of flow betweenness centrality (Brandes, 2001).

### 1.5.2 The swing equation

The swing equation provides a fully dynamic, nonlinear model of power flow. It is based on the notion of modelling each node in the network as a rotating machine. For example, conventional generators possess a rotating core that spins within electrical windings. A detailed treatment of the swing equation is given by Filatrella et al. (2008). Let us now derive the swing equation for

an electrical network with $n$ nodes. During normal operation, each rotating machine should spin at the grid reference frequency $\Omega$. The $i^{\text{th}}$ node's phase angle is then

$$\phi_i(t) = \Omega t + \theta_i(t), \tag{1.30}$$

where $\theta_i(t)$ is its phase difference. Conservation of energy at each node means that the mechanical power and electrical power being transmitted from the rest of the network must balance the dissipative and inertial energy:

$$P_i^{\text{mech}}(t) = P_i^{\text{iner}}(t) + P_i^{\text{diss}}(t) + P_i^{\text{elec}}(t), \tag{1.31}$$

where

$$P_i^{\text{iner}} = \frac{1}{2} J_i \frac{\mathrm{d}}{\mathrm{d}t} \left( \frac{\mathrm{d}\phi_i}{\mathrm{d}t} \right)^2, \tag{1.32}$$

and

$$P_i^{\text{diss}} = D_i \left( \frac{\mathrm{d}\phi_i}{\mathrm{d}t} \right)^2. \tag{1.33}$$

$J_i$ is the moment of inertia of node $i$, and $D_i$ is a damping coefficient. Substituting equation (1.30) into equations (1.32) and (1.33), and making the assumption that $|\dot{\theta}_i| \ll \Omega$ gives

$$P_i^{\text{iner}} = J_i \Omega \frac{\mathrm{d}^2\theta_i}{\mathrm{d}t^2}, \tag{1.34}$$

and

$$P_i^{\text{diss}} = D_i \Omega^2 + 2 D_i \Omega \frac{\mathrm{d}\theta_i}{\mathrm{d}t}. \tag{1.35}$$

As in the DC power flow case considered in section 1.5.1, we shall assume that the reactance of each edge is much larger than the resistance, so that the impedance is dominated by the imaginary terms. We also again assume that the voltage across the network is approximately constant at $V_0$. From equation (1.22), the electrical power $P^{\text{elec}}$ is then given by

$$P_i^{\text{elec}} = \sum_{j=1}^{n} V_0^2 B_{ij} \sin\left(\theta_i - \theta_j\right), \tag{1.36}$$

where $B_{ij}$ is the susceptance of edge $ij$. Equations (1.34), (1.35) and (1.36) can then be substituted into (1.31):

$$J_i\Omega\frac{\mathrm{d}^2\theta_i}{\mathrm{d}t^2} + 2D_i\Omega\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = P_i^{\mathrm{mech}} - D_i\Omega^2 - \sum_{j=1}^{n} V_0^2 B_{ij}\sin\left(\theta_i - \theta_j\right). \quad (1.37)$$

By introducing

$$\gamma_i := \frac{2D_i}{J_i}, \quad (1.38)$$

$$P_i := \frac{P_i^{\mathrm{mech}} - D_i\Omega^2}{J_i\Omega}, \quad (1.39)$$

and

$$K_{ij} := \frac{V_0^2 B_{ij}}{J_i\Omega}, \quad (1.40)$$

equation (1.37) can be written as

$$\frac{\mathrm{d}^2\theta_i}{\mathrm{d}t^2} + \gamma_i\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = P_i - \sum_{j=1}^{n} K_{ij}\sin\left(\theta_i - \theta_j\right). \quad (1.41)$$

Equation (1.41) is the standard form of the swing equation: an $n$-dimensional system of coupled second-order oscillators that specify the power flow in an AC electrical network. The parameter $P_i$ is the effective power injected at node $i$, with units of $\mathrm{s}^{-2}$. If the node is a power source then $P_i > 0$, and $P_i < 0$ if it is a sink. The coefficients $\gamma_i$ and $K_{ij}$ have units of $\mathrm{s}^{-1}$ and $\mathrm{s}^{-2}$ respectively. Typical values of $P_i$, $\gamma_i$ and $K_{ij}$ are all $\mathcal{O}(1)$ for a realistic power grid. For normal grid operation, each node must be synchronised to the grid frequency $\Omega$, which occurs when the condition

$$P_i = \sum_{j=1}^{n} K_{ij}\sin\left(\theta_i - \theta_j\right) \quad (1.42)$$

is satisfied for all $i$. Finding the steady states of a power grid therefore amounts to finding solutions to equation (1.42).

The swing equation is commonly used in electrical engineering (Glover et al., 2016), and allows researchers to investigate the transient stability of power lines (Fouad and Vittal, 1991). It also plays an integral part in popular commercial software used in power systems engineering, such as SPICE. It

should be noted that the power sinks, also sometimes referred to as loads, are not necessarily rotating machines. They may instead be constant, ohmic sinks. In this case, if there are $k$ such ohmic sink nodes, the $n$-dimensional swing equation can be reduced to dimension $n - k$ (Manik et al., 2014). However, this renders the role of the network topology ambiguous; the equations representing the sink nodes will have been effectively absorbed into the equations for the rest of the network. Fortunately, it has been shown that the dynamics of a system containing constant ohmic sinks leads to equations whose form is equivalent to the $n$-dimensional swing equation (Bergen and Hill, 1981). This leads to the structure preserving model and synchronous motor model, and allows any ohmic sink nodes to be effectively treated as rotating machines. Refer to Nishikawa and Motter (2015) for an in-depth discussion of this point. Additionally, some generators such as solar panels may not be rotating machines. However these generators are typically connected into the grid using control mechanisms such as phase-locked-loops (Chung, 2000) that allow them to synchronise to the grid frequency $\Omega$ like a conventional generator. More sophisticated control mechanisms are currently being developed, such as the virtual synchronous machine (Chen et al., 2011), that will further allow renewable sources to approximate conventional rotating machines.

The swing equation has seen increased interest from within the complex systems and network science community. Manik et al. (2014) investigated the stability of a simple two node, one edge swing equation system. One node was chosen to be a source with power $+P$ and the other to be a sink with power $-P$. They showed that there is a saddle node bifurcation at the critical value $2P = K$. Beneath this value no stable synchronous states exist, since there is insufficient coupling $K$ to support the $2P$ units of power being transmitted across the edge. Above this critical value there is a single attracting state representing the stable operating point of the power grid. Menck et al. (2013; 2014) used numerical simulations to analyse the stability of the swing equation posed on large, complex networks. They showed networks containing dead-ends or tree like substructures were particularly unstable. Rohden et al. (2012) showed that increasing the numbers of generators on a network can sometimes decrease power grid stability. Witthaut and Timme

(2012) observed that the addition of new edges into the swing equation can cause a loss of stability, a phenomenon they relate to Braess's paradox as discussed in section 1.3.

## 1.6 Summary

This chapter has introduced the main themes and motivation of the thesis, and reviewed the relevant literature. Aspects from each of the reviewed topics will be drawn upon throughout the thesis in aid of the overall objective: to analyse the efficiency and resilience of networks with heterogeneous node behaviour. Network science will underpin much of the work, with nonlinear dynamical models as well as steady state models being used to evaluate network flows. Additionally, the Price of Anarchy will be used to assess flow efficiency.

Several deficiencies are apparent from the literature, that this thesis will seek to address. Firstly, neither the Price of Anarchy nor network resilience to cascades has been analysed as a function of the composition and behaviour of network nodes. Secondly, cascading failures in networks tend to be investigated using steady state models rather than nonlinear dynamical models. And thirdly, the field of network science has paid little attention to electrical "micro-grids"; small-scale, highly dynamic electricity grids containing high volumes of renewable sources. Each of these issues will be addressed in the following chapters.

# Chapter 2

# Redundancy and the Price of Anarchy in flow networks

## 2.1 Introduction

Flows on networks, such as traffic taking routes of shortest travel time or electrical current taking paths of least resistance though a network, can waste resources because they follow a local rather than a system-wide optimisation of the flow (Cohen and Horowitz, 1991; Youn et al., 2008; Çolak et al., 2016). In the case of traffic this is because drivers choose their routes non-cooperatively or "selfishly", whilst for electricity the behaviour stems from physical conservation laws. Nevertheless, the efficiency of these systems can often be improved, either by attempting to alter the flow dynamics, limiting flow volume, or changing the underlying network structure. Understanding the relationship between efficiency, flow volume and network structure is therefore important for any system operator attempting to optimise the flow. Additionally, flow inefficiency may be caused in part by structural redundancy in the network, which might in itself be desirable for purposes of resilience and robustness. Such a situation would present the operator with a trade-off to be made between efficiency and resilience.

This chapter gauges this flow inefficiency using the Price of Anarchy $\mathcal{P}$

and investigates its relationship with structural redundancy, particularly for physically constrained flows such as electrical current. The chapter then focuses on how $\mathcal{P}$ depends upon network structure and specifically upon the behaviour and composition of the network's nodes. This is motivated by the increasing proliferation of distributed renewable sources of energy in modern electrical grids. This paradigm shift means modern power networks may no longer contain only a few very large sources, but many small generators embedded throughout their structure. The numbers of generators may even vary throughout a day, influenced by both external conditions and network controllers. Uncovering the dependence of $\mathcal{P}$ on the proportions of generators and consumers is therefore an important step to understanding how distributed generation impacts efficiency and redundancy.

The main focus of this chapter is therefore on the relationship between $\mathcal{P}$ and redundancy, and its dependence on node behaviour. The work presented in this chapter was published in Smith et al. (2019). Section 3.2 defines $\mathcal{P}$ for flow networks and details how the relevant quantities are computed, as well as demonstrating the dependence of $\mathcal{P}$ on network structure. Section 3.3 establishes the relationship between non-cooperative "selfish" behaviour and least-resistance flows, such as electrical current, and defines a measure of network redundancy. The relationship between $\mathcal{P}$ and redundancy is then demonstrated on a variety of networks; both synthetic random networks and real power grid topologies. Section 3.4 then investigates the dependence of $\mathcal{P}$ on network node behaviour.

## 2.2 The Price of Anarchy $\mathcal{P}$ in flow networks

### 2.2.1 Definition and computation of $\mathcal{P}$

This section introduces the network flow model and the formulation of $\mathcal{P}$ to be used throughout the chapter. It is broadly the same as the model introduced by Roughgarden (2003), but with variable and heterogeneous source and sink terms. All flows are modelled as moving though networks of the form

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $n = |\mathcal{V}|$ nodes and $m = |\mathcal{E}|$ edges. $n_+$ nodes are chosen to be sources of flow, $n_-$ to be sinks, and the remaining $n_p$ nodes are passive or empty. All networks satisfy the condition $n_+ + n_- + n_p = n$. The flows are composed of individual units incurring a cost each time they traverse an edge. The network edges may represent represent roads or electrical transmission lines depending on the context. The individual units of flow comprise only a very small amount of the overall flow volume and so, following Roughgarden (2005), the flow is approximated as a continuous quantity. This is the so-called non-atomic flow approximation and is commonly employed in traffic flow modelling (e.g. Coclite et al., 2005) and further work on the non-cooperative flow (e.g. Youn et al., 2008). Each edge $e \in \mathcal{E}$ then has a continuous flow volume $f_e \in \mathbb{R}$, with its cost of traversal given by a linear function

$$c_e(f_e) = \alpha_e f_e + \beta_e. \tag{2.1}$$

The functions $c_e$ represent travel time in the context of a traffic flow. For an electric current flow, the functions represent voltage drop with the coefficients $\alpha_e$ and $\beta_e$ denoting Ohmic resistance and flow independent voltage drops respectively. Electrical networks form the main motivation for this work and so we shall constrain our attention to linear cost functions, since network voltages are typically linear functions of current. The total cost borne across all infinitesimal units on an edge $e$ is $c_e(f_e)f_e$, representing power loss in the electrical context. The total cost across the entire network is

$$\mathcal{C}(f) = \sum_{e \in \mathcal{E}} c_e(f_e) f_e, \tag{2.2}$$

where $f \in \mathbb{R}^m$ is the flow vector. The global optimal flow $f_{\mathrm{GO}}$ is the flow pattern that minimises the total cost:

$$\min_f \ \mathcal{C}(f) \quad \text{constrained by } Ef = b, \tag{2.3}$$

where $E \in \mathbb{R}^{n \times m}$ is the node-edge incidence matrix and $b$ is the flow injection vector with components

$$b_v = \begin{cases} (1+\xi_v)F/n_+, & \text{if node } v \text{ is a source,} \\ -(1+\xi_v)F/n_-, & \text{if node } v \text{ is a sink,} \\ 0, & \text{otherwise,} \end{cases} \tag{2.4}$$

with $F$ being the total flow or current injected into the network. $\xi_v$ is a normally distributed random variable with mean 0, and is used to induce heterogeneity in the strength of the sources and sinks. The variance of $\xi_v$ will be chosen small enough to be unlikely to cause the source-sink terms in equation (2.4) to change sign. In the rare event that a sign change does occur, the result will be discarded and $\xi_v$ drawn again. The condition $Ef = b$ enforces conservation of flow at nodes, also known as Kirchoff's current law.

Note that the source-sink strength heterogeneity can effectively be turned off by setting $\xi_v = 0 \; \forall v$. This gives the case where all sources output the same volume of flow and all sinks consume the same volume of flow. This homogeneous source-sink behaviour will used throughout most of this chapter. Unless explicitly stated otherwise, the reader is to assume that $\xi_v = 0 \; \forall v$ in the experiments shown in subsequent sections.

Recall that a Nash equilibrium flow $f_{\text{Nash}}$ is a state in which none of the flow units can decrease their cost. Equivalently, it is a state where the cost of all routes between an arbitrarily chosen source-sink pair is the same. This is the equilibrium reached when the flow acts non-cooperatively or "selfishly" and is given by the optimisation problem

$$\min_f \sum_e \int_0^{fe} c_e(q)\,\mathrm{d}q \quad \text{constrained by } Ef = b. \tag{2.5}$$

Refer to Roughgarden (2005) for a full derivation of equation (2.5). Both (2.3) and (2.5) are convex, quadratic objective functions that are solved using standard optimisation methods, as detailed in Appendix I. The Price of

Anarchy $\mathcal{P}$ in the system is then defined as

$$\mathcal{P} := \frac{\mathcal{C}(f_{\text{Nash}})}{\mathcal{C}(f_{\text{GO}})} \equiv \frac{\mathcal{C}_{\text{Nash}}}{\mathcal{C}_{\text{GO}}}, \tag{2.6}$$

being cost of the Nash flow relative to the cost of the global optimum. It will be shown in Section 2.3 that the physical constraints of electrical flows are equivalent to Nash equilibria, and $\mathcal{P}$ can therefore be used to probe the inherent network inefficiency of electrical networks.

### 2.2.2 Examples in lattices

Since the edge cost functions in equation (2.1) are linear, $\mathcal{P}$ is bounded from above by a value of $4/3$ (Roughgarden, 2003). However, away from this upper limit the values of $\mathcal{P}$ are strongly influenced by flow volume $F$, network topology, and the proportions of flow source and sink nodes. The dependence on flow volume and topology has been examined in Youn et al. (2008); Skinner (2015); Rose et al. (2016); however, the dependence upon source and sink proportions has not been previously investigated. This dependence on node behaviour is the focus of Section 2.4.

To demonstrate the impact of node behaviour, and provide motivation for following sections, it is instructive to show how $\mathcal{P}$ varies with flow volume in some simple example networks. Let us start though with the Pigou network (Pigou, 1920) as introduced in Chapter 2 and depicted in figure 2.1(a). This network will also be useful in the next section for demonstrating the relationship between $\mathcal{P}$ and network redundancy. Being the simplest network with a value of $\mathcal{P} > 1$, the Pigou network admits an analytic description of $\mathcal{P}$ as a function of $F$, which will now be derived. The Pigou network consists of two nodes connected by two edges. One node is a flow source and the other a sink, generating and consuming respectively a volume $F$. The cost on the upper edge, labelled edge 1, is equal to the flow volume $f_1$. The cost on edge 2 is equal to unity and is independent of the flow volume $f_2$. In the notation of equation (2.1), the coefficients on the edges are thus $(\alpha_1, \beta_1) = (1, 0)$ and $(\alpha_2, \beta_2) = (0, 1)$. Each infinitesimal unit of flow takes either the variable

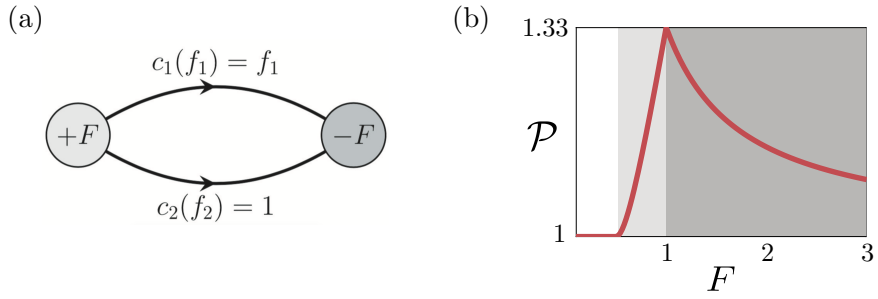



Figure 2.1: (a) The Pigou network. The left and right nodes are a source and sink respectively. A total flow volume of $F$ flows through the network. The upper edge, labelled edge 1, has a variable cost function. Units of flow traversing edge 1 therefore incur a cost equal to the flow volume on the edge. Edge 2 has a fixed cost, independent of the flow volume. (b) The Price of Anarchy $\mathcal{P}$ as a function of $F$ in the network, which has been computed numerically from the optimisation problems (2.3) and (2.5) using the algorithm outlined in Appendix 1. An analytic description of $\mathcal{P}(F)$ is given in equation (2.7).

cost edge 1, or the fixed cost edge 2 to travel from the source to the sink.

At the Nash equilibrium, which is reached when individual flow units act un-cooperatively, the flow will mostly tend to use the cheaper, variably costed edge 1. This will lead to congestion and inefficiency. Optimally, some flow should take the more expensive edge 2. Recall also that at the Nash equilibrium, there is no possible reduction in cost available to the flow units by changing route. Figure 2.1(b) shows $\mathcal{P}$, which gauges the inefficiency of the Nash equilibrium, as a function of $F$ and illustrates the strong dependence of $\mathcal{P}$ on flow volume. The maximum value of $4/3$ is attained only at a value of $F = 1$, and is typically much less elsewhere. This behaviour can be quantified analytically as follows. For $0 < F \leq 1/2$, indicated by the unshaded area in figure 2.1(b), all flow is routed over edge 1 in both the Nash and GO equilibria, with identical costs $\mathcal{C} = F^2$; consequently $\mathcal{P} = 1$. For $1/2 < F \leq 1$ (light-grey area), $f_1 = F$ under the Nash flow, and so $\mathcal{C}_{\text{Nash}} = F^2$. The GO minimises its cost when $f_1 = 1/2$ and $f_2 = F - 1/2$, giving $\mathcal{C}_{\text{GO}} = F - 1/4$ and so $\mathcal{P} = F^2/(F - 1/4)$. For $F > 1$ (dark-grey area), the Nash equilibrium routes all flow surplus of $F = 1$ through edge 2, giving $\mathcal{C}_{\text{Nash}} = F$, whereas the GO remains unchanged from the previous

regime, hence $\mathcal{P} = F/(F - 1/4)$. In summary, for the Pigou network:

$$\mathcal{P} = \begin{cases} 1, & \text{for } 0 < F \leq 1/2, \\ F^2/(F - 1/4), & \text{for } 1/2 < F \leq 1, \\ F/(F - 1/4), & \text{for } F > 1. \end{cases} \tag{2.7}$$

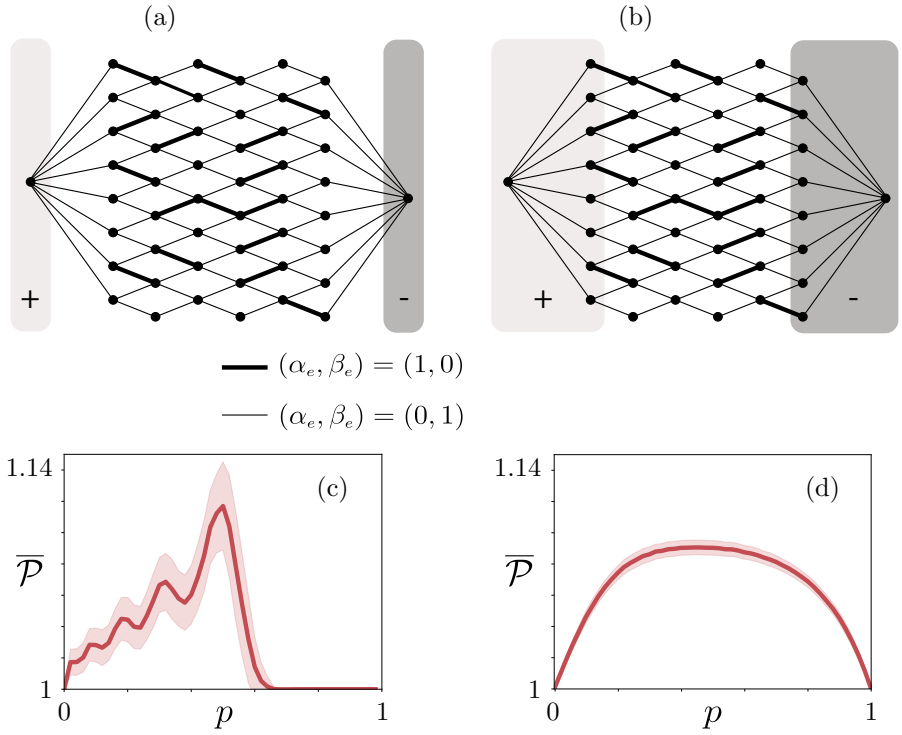



Figure 2.2: (a) and (b) show instances of lattices, with thick and thin lines indicating edges with variable and fixed cost respectively. Edges are selected uniformly at random to be of the variable type, according to a probability $p$. The lattices in this example have a height of 8 nodes and a width of 6. In both (a) and (b) the lattices are shown in an instance where $p = 0.25$ so that approximately a quarter of edges should therefore be chosen to have a variable cost. Nodes sitting within the light-grey regions are sources of flow, while those within the dark-grey region are sinks. Panels (c) and (d) show $\overline{\mathcal{P}}$ as a function of $p$ for an ensemble of 2000 lattices. (c) is the case for a single source-sink pair as illustrated in (a), while (d) is the many source-sink case as illustrated in (b). Red shaded areas indicate standard deviation.

From the above example, it is clear that $\mathcal{P}$ is highly dependent on $F$. However, network topology and the distribution of edge cost parameters also have a strong influence. This was demonstrated in Rose et al. (2016) using flow networks with a lattice topology, as depicted in figure 2.2(a). It was shown that $\mathcal{P}$ predictably depends upon the fraction of variably costed edges, as well as the size and aspect ratio of the lattice. These results of Rose et al. will now be recreated and compared to those obtained from a modified lattice model, to show that the proportion of source and sink nodes in a network is also an important factor.

The lattice networks under consideration have a random mixture of variable and fixed cost edges. In particular, each edge is chosen with a probability $p$ to have variable cost, with coefficients $(\alpha_e, \beta_e) = (1, 0)$. Remaining edges have a fixed cost with coefficients $(\alpha_e, \beta_e) = (0, 1)$. A lattice with $p = 0$ will therefore contain only edges whose costs are constant and independent of flow volume. A lattice with $p = 1$ will be completely variable; all of its edges will be a linear function of flow volume. The mean Price of Anarchy $\overline{\mathcal{P}}$ is then computed as a function of $p$ over an ensemble of 2000 lattice network realisations.

The lattice networks considered in Rose et al. (2016), an example of which is depicted in figure 2.2(a), had only one source and one sink node. The resulting $\overline{\mathcal{P}}$ profile, shown in figure 2.2(c), undergoes a series of ripples as $p$ is increased from 0 to 1. These ripples were identified and explained by Rose et al. (2016) as being equal in number to the number of nodes across the width of the lattices. It is now shown that these ripples can be suppressed by introducing more source and sink nodes into the lattice. Figure 2.2(b) shows a lattice with many source and sink nodes; each node in the left-most column of the lattice is a source, and each node in the right-most column is a sink. The resulting $\overline{\mathcal{P}}$ profile in figure 2.2(d) is a flattened, rounded curve. $\overline{\mathcal{P}}$ is now more broadly spread across the range of $p$, with a diminished peak. Furthermore, the results are less noisy than those in figure 2.2(c), as indicated by the significantly reduced standard deviation. This flattening and de-noising of the $\overline{\mathcal{P}}$ curve can be ascribed to the system being more tightly restricted by the increased number of sources and sinks. More source and

sink nodes, and therefore fewer passive or empty nodes, constrains the space of available flow patterns, thus limiting the possible difference between the Nash and the global optimum. In general, $\overline{\mathcal{P}}$ decreases as passive nodes are replaced with sources and sinks. Further examples of this will be shown in Section 2.4, which investigates the impact of source-sink numbers on general complex networks.

In summary, this section has introduced the network model for computing $\mathcal{P}$ in flow networks and provided preliminary evidence that $\mathcal{P}$ is strongly influenced by topology, edge cost functions, flow volume and, crucially, source-sink numbers. The rest of this chapter uses this model to investigate the role of source-sink numbers and network topology, and interprets the results for physical flows such as electricity.

## 2.3 Relationship with redundancy

In the literature, it is conventionally assumed that flows tend towards the Nash equilibrium through "selfish" or un-coordinated behaviour, for instance in Roughgarden (2003; 2005); Youn et al. (2008). This allows the possibility that the Nash flow can in principle move towards the global optimum through some form of central management or enforced cooperation. This is true in traffic flows where drivers can be pushed towards more optimal behaviour through in-car traffic routing apps and adaptive road speed limits, as explored in Çolak et al. (2016) for example. In some systems, however, the global optimum may not be available since the system is constrained by physical conservation laws to behave as a Nash equilibrium. This section reveals that this is the case for electrical flows. The section then interprets the meaning of the global optimum in an electrical flow context, showing that it is related to a novel measure of topological redundancy.

### 2.3.1 Nash equilibria and Kirchoff's laws

Recall that the Nash equilibrium is given by the optimisation problem in equation (2.5), and is the state in which no reduction in cost is available to a unit of flow by changing route. Equivalently, this means the cost of every route between an arbitrary source-sink pair must be equal. The cost functions can be interpreted as voltages in an electrical flow context. The Nash flow must also satisfy conservation of flow, also known as Kirchoff's current law (KCL). A physical interpretation of the Nash equilibrium now comes from a consideration of Kirchoff's voltage law (KVL), which states that voltages around closed cycles in an electrical network sum to zero (Alexander and Sadiku, 2012). If there is a cycle embedded in a network, then there will be at least two distinct paths between a pair of source and sink nodes. At the Nash equilibrium, each arm of the cycle must have equal cost; hence the cost of any traversal around the cycle is zero, and so the Nash equilibrium condition is equivalent to KVL. The Nash flow therefore necessarily satisfies both Kirchoff's current and voltage laws and is thus a physically legitimate electrical flow for an electrical network in stable operation with matched supply and demand. The relative inefficiency of this flow, resulting in $\mathcal{P} > 1$, thus stems from the constraints of Kirchoff's conservation laws that define the Nash equilibrium.

### 2.3.2 Defining a redundancy measure, $\mathcal{R}$

$\mathcal{P}$ measures the disparity between the costs associated with the Nash and GO flows. In an electrical context the GO would correspond to a flow being able to violate KVL in order to minimise total power loss; however, such an equilibrium would nevertheless be desirable to obtain because it minimises the power consumed by the network. Therefore, $\mathcal{P}$ remains a useful metric for assessing efficiency in networks with flows following paths of least resistance, and also for topological redundancy as we now show.

Consider again the Pigou network from figure 2.1(a), whose values of $\mathcal{P}$ as a function of $F$ are given in equation (2.7). The Nash equilibrium, equivalent

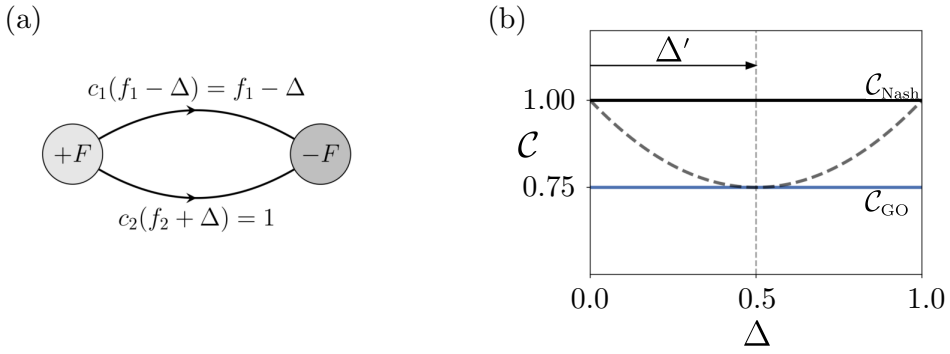



Figure 2.3: (a) The Pigou network. A flow volume $\Delta$ can be transfered from edge 1 to edge 2 by capping the flow capacity of edge 1. (b) The dashed line shows how the cost of the Nash flow changes with increasing $\Delta$ for $F = 1$, as given by equation (2.11). The cost of the Nash and GO flows on the unmodified network are indicated by the solid lines. At $\Delta' = 1/2$, the Nash flow is optimised and coincides with the GO.

to a flow obeying both KVL and KCL, and whose cost is $\mathcal{C}_{\text{Nash}}$, can be driven towards the GO by manipulating the network such that excess flow is transferred from edge 1 to edge 2. Specifically, an amount $\Delta$ of the flow from edge 1 in the network can be removed by capping the edge's capacity in order to reduce the cost of the Nash flow, as shown in figure 2.3(a). Figure 2.3(b) shows how the cost of this modified Nash flow decreases with $\Delta$ for a total flow volume in the network of $F = 1$. The Pigou network admits a simple analytic expression for the cost of this modified Nash flow:

$$\mathcal{C}_{\text{Nash}}(F, \Delta) = (f_1 - \Delta)^2 + f_2 + \Delta, \tag{2.8}$$

where $f_1$ and $f_2$ are the flow volumes on edge 1 and edge 2 respectively at the Nash equilibrium, when the total flow volume is $F$. $f_1$ and $f_2$ are given for any value of $F$ in section 2.2.2. Using $f_1 = F - f_2$, equation (2.8) becomes

$$\mathcal{C}_{\text{Nash}}(F, \Delta) = (F - f_2 - \Delta)^2 + f_2 + \Delta. \tag{2.9}$$

To find the optimum reduction $\Delta'$ that minimises $\mathcal{C}_{\text{Nash}}$, we need to find $\Delta$ such that

$$\frac{\partial \mathcal{C}_{\text{Nash}}}{\partial \Delta} = 0. \tag{2.10}$$

Recall from section 2.2.2 that at $F = 1$ all of the flow uses edge 1 under Nash conditions. Consequently, for $F = 1$, we have $f_1 = 1$ and $f_2 = 0$. Equation (2.9) then becomes

$$\mathcal{C}_{\text{Nash}}(F, \Delta) = (F - \Delta)^2 + \Delta, \tag{2.11}$$

and so

$$\frac{\partial \mathcal{C}_{\text{Nash}}}{\partial \Delta} = 2\Delta - 2F + 1. \tag{2.12}$$

Equating the above to zero reveals that the optimum is $\Delta' = F - 1/2 = 1/2$. Indeed, for any $F \leq 1$, all of the flow takes edge 1 under Nash conditions. Thus $f_1 = F$ and $f_2 = 0$ for $F \leq 1$. It is easy to check that $\Delta' = F - 1/2$ therefore holds for $1/2 \leq F \leq 1$. Note that $\Delta' = F - 1/2$ is negative for $F < 1/2$. This is not physical and means that no reduction in cost is possible for $F < 1/2$. From section 2.2.2, we know that for $F > 1$, the Nash equilibrium is $f_1 = 1$ and $f_2 = F - 1$. Consequently,

$$\mathcal{C}_{\text{Nash}}(F, \Delta) = (1 - \Delta)^2 + F - 1 + \Delta, \quad \text{for } F > 1. \tag{2.13}$$

This gives

$$\frac{\partial \mathcal{C}_{\text{Nash}}}{\partial \Delta} = 2\Delta - 1, \tag{2.14}$$

and so $\Delta' = 1/2$. In summary, $\Delta' = F - 1/2$ for $1/2 \leq F \leq 1$ and $\Delta' = 1/2$ for $F \geq 1$, with no reduction in cost available for $F \leq 1/2$.

For notational brevity we define $\mathcal{C}_{\text{Nash}} = \mathcal{C}_{\text{Nash}}(F, 0)$ and $\mathcal{C}'_{\text{Nash}} = \mathcal{C}_{\text{Nash}}(F, \Delta')$. The equilibrium on the modified network, by definition, has cost $\mathcal{C}'_{\text{Nash}} \leq \mathcal{C}_{\text{Nash}}$. This means that edge 1 provides redundant capacity that can be removed. Defining this edge redundancy in terms of the costs gives:

$$\mathcal{R}^1 := \frac{\mathcal{C}_{\text{Nash}} - \mathcal{C}'_{\text{Nash}}}{\mathcal{C}_{\text{Nash}}} = \begin{cases} 0, & 0 \leq F < 1/2, \\ \left(F - 1/2\right)^2 / F^2, & 1/2 \leq F < 1, \\ 1/(4F), & F > 1, \end{cases} \tag{2.15}$$

which is the relative decrease in cost available by removing capacity from edge 1. Note that this redundancy is flow volume dependent; it varies as a function of $F$. The capacity reduction $\Delta'$ enforces the flow volume on edge 1

to be equal to the volume on edge 1 under GO flow conditions. The GO flow pattern therefore reveals the optimal capacity reduction $\Delta'$, as illustrated in figure 2.3(b). In summary, a capacity cap $\Delta$ should be enforced to improve the efficieny of the Nash flow. The optimal capacity cap $\Delta'$ is given by the difference between the flow volumes on edge 1 for each equilibria, i.e., $\Delta' = f_1^{\text{Nash}} - f_1^{\text{GO}}$.

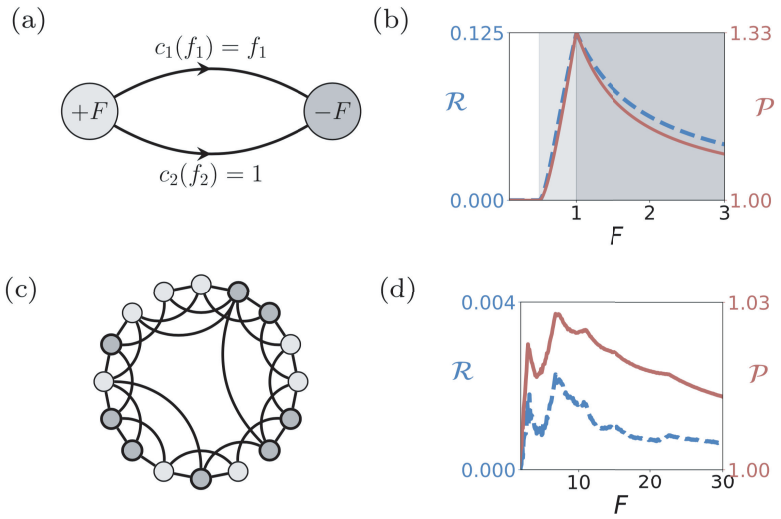



Figure 2.4: (a) The example Pigou network. (b) $\mathcal{P}$ (red line) and $\mathcal{R}$ (blue dashed line) for the Pigou network in (a) are shown as functions of $F$. (c) A small world network with $q = 0.1$, $k = 4$, $n = 16$ and $(n_+, n_-, n_p) = (8, 8, 0)$. (d) $\mathcal{P}$ and $\mathcal{R}$ as functions of $F$ for the small world network shown in (c).

In order to generalise the redundancy measure in equation (2.15) to larger networks it is averaged over both edges to give $\mathcal{R} := \overline{\mathcal{R}^e}$. The measure $\mathcal{R}$ is therefore the mean decrease in cost attainable by removing capacity from an edge. Figure 2.4(b) shows $\mathcal{R}$ as a function of $F$ in the Pigou network. The figure also shows $\mathcal{P}$ as a function of $F$ to demonstrate the correspondence between the two measures. Their forms emulate each other, with peaks occurring at the same values of $F$. Therefore, despite an electrical flow being constrained by physics to operate at the Nash equilibrium, computing $\mathcal{P}$ and $f_{\text{GO}}$ is still useful since $\mathcal{P}$ acts as a proxy for edge redundancy and $f_{\text{GO}}$ gives the volume of redundancy on each edge. A high value of $\mathcal{P}$ therefore indicates that a network's edges are providing excess capacity; reducing this capacity will reduce power loss over the network.

### 2.3.3 Computing $\mathcal{R}$ in complex networks

For larger networks with overlapping paths, there is unlikely to be an analytic expression for $\Delta'$ and so Algorithm 1 is used. This algorithm, summarised in the pseudo-code listing below, uses knowledge of the GO flow pattern to evaluate the optimal amount by which to cap each edge. The GO and Nash flows are computed via the optimisation problems (2.3) and (2.5) respectively. The algorithm begins by computing the Nash and GO flow on the given network. It then iterates over the edges, and on each edge checks whether the Nash flow volume is greater than the GO flow volume. If it is, then an upper volume cap equal to the GO flow volume is placed on that edge and the Nash flow recomputed. This adjusted Nash flow will be more efficient and have a lower cost, since redundant capacity has now been effectively removed by the cap. The relative decrease in cost is then recorded. The mean of these decreases then gives the network's redundancy measure $\mathcal{R}$.

**Algorithm 1** Compute $\mathcal{R}$

**Input:** A network $\mathcal{G} = (\mathcal{V}, \mathcal{E}, c)$
**Output:** The redundancy measure $\mathcal{R}$ on $\mathcal{G}$

Compute the Nash and GO flows $f_{\text{Nash}}$ and $f_{\text{GO}}$
$\mathcal{C}_{\text{Nash}} = \sum_{e\in\mathcal{E}} c_e(f^e_{\text{Nash}}) f^e_{\text{Nash}}$
**for** $e \in \mathcal{E}$ **do**
    **if** $f^e_{\text{Nash}} > f^e_{\text{GO}}$ **then**
        Set an upper limit $\kappa = f^e_{\text{GO}}$ on edge $e$
        Compute modified Nash flow $f'_{\text{Nash}}$
        $\mathcal{C}'_{\text{Nash}} = \sum_{e\in\mathcal{E}} c_e(f'^e_{\text{Nash}}) f'^e_{\text{Nash}}$
        $\mathcal{R}^e = (\mathcal{C}_{\text{Nash}} - \mathcal{C}'_{\text{Nash}})/\mathcal{C}_{\text{Nash}}$
    **else**
        $\mathcal{R}^e = 0$
    **end if**
**end for**
$\mathcal{R} = \overline{\mathcal{R}^e}$

### 2.3.4 Correspondence between $\mathcal{R}$ and $\mathcal{P}$

Using Algorithm 1, the redundancy measure $\mathcal{R}$ can now be computed for complex networks such as the small-world network depicted in figure 2.4(c).

This network has 16 nodes, a mean degree of 4 and a random rewiring parameter of $q = 0.1$. It has equal numbers of sources and sinks but no passive nodes; i.e., $(n_+, n_-, n_p) = (8, 8, 0)$. Figure 2.4(d) shows $\mathcal{R}$ as a function of flow volume $F$ in this network. It exhibits a series of peaks, attaining a maximum at $F \approx 8$. Plotted in the same figure is the Price of Anarchy $\mathcal{P}$, which displays peaks in the same places as $\mathcal{R}$. This similarity in the behaviour of $\mathcal{P}$ and $\mathcal{R}$ was also observed in the Pigou network in figure 2.4(b).

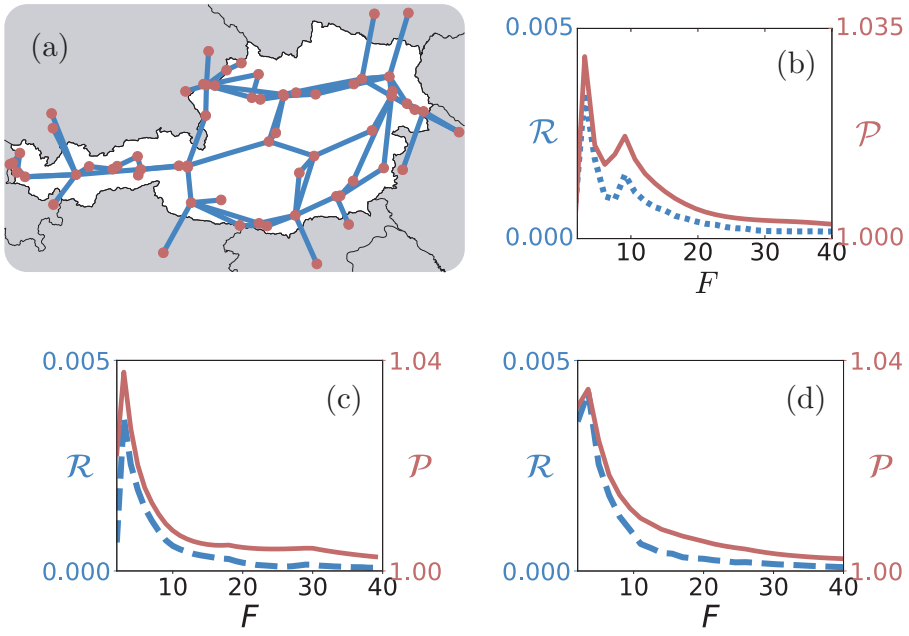



Figure 2.5: (a) The Austrian power grid, constructed from open source topological data from (APG, 2017). (b) $\mathcal{P}$ and $\mathcal{R}$ for the network in (a) as a function of total current $F$, where $F$ has been normalised using the per-unit system. Panels (c) and (d) show $\mathcal{P}$ and $\mathcal{R}$ for the IEEE 14 bus and 118 bus test networks respectively, where the flow has again been normalised into the per-unit system.

To further examine the correspondence between $\mathcal{R}$ and $\mathcal{P}$ they will now be computed for several real world networks, starting with the Austrian national power grid plotted in figure 2.5(a). Obtained from APG (2017), the Austrian data set contains $n = 67$ nodes and $m = 85$ edges. Each node is a substation, behind which may be either a power station or a distribution grid which is unresolved by the data. Generators and distribution nodes

represent sources and sink respectively. The edges are power transmission lines. 11 nodes are generators, and the remainder are distribution nodes so that $(n_+, n_-, n_p) = (11, 56, 0)$. The data set contains only topological information; power output and consumption data is not available. Therefore, power input and output is modelled as per equation (2.4) with $\xi_v = 0$ so that sources generate an equal amount, sinks consume an equal amount, and flow supply-demand is matched. Heterogeneous levels of generation and real world power consumption data will be treated in the next section and in Chapter 7 respectively; the objective here however is simply to demonstrate the dependence of $\mathcal{R}$ as a function of $F$ on a realistic topology. Figure 2.5(b) shows the $\mathcal{R}$ versus $F$ profile for the Austrian grid. $\mathcal{R}$ peaks at two values of $F$ before decreasing 0 for high $F$, indicating that for very high flow volumes there is little redundancy since all edges are heavily used. $\mathcal{P}$ is also plotted alongside $\mathcal{R}$, once again demonstrating the similarity between the two measures. The resistance and voltage drop coefficients for the power loss functions in the Austrian grid are unknown from the data and so, following Youn et al. (2008), they are modelled here as uniform random variables in the range $[0, 1]$. As noted in Youn et al. (2008), rescaling these variables simply gives a 1:1 rescaling in the $F$ axis. The $F$ axis can therefore be regarded as the flow volume relative to the edge coefficients; this type of normalisation is common in electrical engineering where it is referred to as the per-unit system (Alexander and Sadiku, 2012).

Next, $\mathcal{R}$ and $\mathcal{P}$ are computed as a function of $F$ for the well-known IEEE test networks; specifically, the 14 bus and 118 bus test networks. These networks are widely used in the electrical engineering literature as test beds for new algorithms and models. They are snapshots of subsets of the United States power transmission grid from the 1960s. Nodes and edges have the same meaning as in the Austrian grid data. The 14 bus network has $n = 14$ nodes and $m = 20$ edges. The 118 bus network has $n = 118$ nodes and $m = 179$ edges. All other quantities are modelled in the same way as in the Austrian case above. The resulting $\mathcal{R}$ and $\mathcal{P}$ profiles are displayed in figures 2.5(c) and (d). Here the peak values of $\mathcal{P}$ are $\sim 1.035$, corresponding to a value of $\mathcal{R}$ indicating an average 0.4% increase in efficiency available to the whole system from reducing the capacity of a single edge. As this is a per

edge value, it reveals a substantial amount of inefficiency across the network as a whole.

Key to what follows in the next section is that the maximum values of $\mathcal{P}$ and $\mathcal{R}$ occur at the same flow volume $F$. Determination of $\mathcal{R}$ is computationally onerous, requiring the evaluation of a convex optimisation problem for each of a network's edges, rendering it impractical for all but the smallest of networks. This is because the Nash flow must be recomputed after capping each edge. Evaluating $\mathcal{P}$ therefore provides a simple computational proxy for identifying regimes of relative redundancy, enabling very large networks of complex topology and composition to be investigated. The next section will investigate how $\mathcal{P}$ depends upon the network's node composition, revealing how regimes of high $\mathcal{P}$ and therefore redundancy are influenced by the proportions of source and sink nodes.

## 2.4 Dependence of $\mathcal{P}$ on node composition and function

The previous two sections have defined $\mathcal{P}$ in the context of flow networks and established its use as a proxy for redundancy in the case of physically constrained flows such as electricity. This section investigates how $\mathcal{P}$ depends on network composition, and particularly on source-sink node proportions.

We first consider networks whose source and sink nodes have equal flow outputs and inputs respectively, given by the case where $\xi_v = 0$ for all $v \in \mathcal{V}$ in equation (2.4). For a total flow volume $F$, the dependencies of $\mathcal{P}$ on network structure and composition are obtained from an ensemble of 1000 such random small-world network realisations, constructed using the Watts–Strogatz method (Watts and Strogatz, 1998). These networks are parameterised by the rewiring probability $q \in [0, 1]$, initial degree $k$, and the number of nodes $n$. A network with $q = 0$ is a regular lattice, while a network with $q = 1$ is a completely disordered random network. As usual, a given network's $n$ nodes comprise $n_+$, $n_-$ and $n_p$ source, sink and passive nodes, respectively. In this section, the location of these nodes will be allocated uniformly at random. The edge cost coefficients $\alpha_e$ and $\beta_e$ are both uniformly

distributed random variables in the range $[0, 1]$.

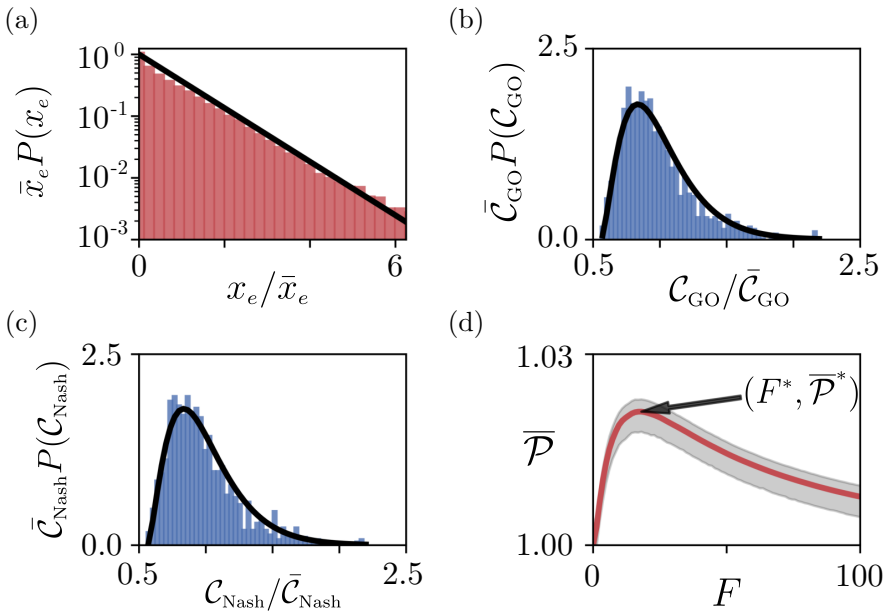



Figure 2.6: (a) The Nash equilibrium edge power $x_e = c_e(f_e)f_e$ distribution in small world networks with $q = 0.1$, $k = 4$, $n = 32$, $n_+ = n_- = 16$, $n_p = 0$. Distributions of (b) $\mathcal{C}_{\rm GO}$ and (c) $\mathcal{C}_{\rm Nash}$ in an ensemble of 1000 such networks with total flow volume $F = 20$. The solid lines are fitted shifted gamma distributions with shape parameter $\nu = 3.44$, scale parameter $\mu = 3.50$ and shift parameter $\sigma = 12.3$ in (b) and $\nu = 3.51$, $\mu = 3.51$ and $\sigma = 12.4$ in (c). The mean of these distributions is given by $\nu\mu + \sigma$. Therefore the mean cost is 24.34 in (b) and 24.72 in (c), indicating that the Nash flow is on average slightly more expensive than the GO for this value of $F$. Both the Nash and the GO distributions have a similar degree of skewness $2/\sqrt{\nu}$. (d) The mean Price of Anarchy $\overline{\mathcal{P}}$ as a function of $F$, with maximum at $(F^*, \overline{\mathcal{P}}^*)$. The shaded region indicates the 95% confidence interval, computed using the statistical bootstrapping method (Efron and Tibshirani, 1986).

The resulting Nash and GO flows on these random networks are found to have exponentially distributed individual edge costs $x_e = c_e(f_e)f_e$, as seen in figure 2.6(a). In the context of electrical current, the value $x_e$ is the power loss over edge $e$. In a traffic context $x_e$ is the total travel time borne by all drivers on the edge. At the macroscopic scale the Nash and GO flow costs

are gamma-distributed with a probability density function

$$P(\mathcal{C}) = \left(\mathcal{C} - \frac{\sigma}{\mu}\right)^{\nu-1} \frac{e^{-\frac{\mathcal{C}-\sigma}{\mu}}}{\mu\Gamma(\nu)}, \tag{2.16}$$

where $\nu$, $\mu$ and $\sigma$ are the shape, scale and shift parameters respectively. This is shown in figures 2.6(b) for the GO flow, and 2.6(c) for the Nash flow, and confirmed with Kolmogorov–Smirnov tests (see Frank and Massey, 1951, for an overview). The emergence of gamma distributed total costs can be attributed to the fact that they are formed from ensembles of exponentially distributed individual edge costs.

Note that the use of a shift parameter $\sigma$ is not standard in a gamma distribution. However, it is necessary here due to the nature of the data. To see why this is, consider that the data in figures 2.6(b) and (c) are the flow costs across an ensemble of randomly generated network topologies connecting a set number of source and sink nodes, with a fixed flow volume $F$. The lowest cost in each ensemble will therefore belong to the most efficient topology in the ensemble. However, for any $F > 0$, even this most efficient topology must result in a non-zero cost. This is because the flow $F$ must still be routed from the sources to the sinks. For example, a volume of traffic will still experience a non-zero total travel time between their origin and destination, no matter how efficient the road network. These lowest costs in the ensemble increase as $F$ increases, shifting the resulting distributions of costs to higher values. The shift parameter $\sigma$ should therefore be regarded as taking into account the total flow volume $F$.

For each value of total flow $F$, the mean of the resulting distribution of $\mathcal{P}$, denoted $\overline{\mathcal{P}}$, is shown in figure 2.6(d). With increasing flow, $\overline{\mathcal{P}}$ rapidly rises to a maximum $\overline{\mathcal{P}}^*$ at $F^*$, before declining to unity. How the values of $\overline{\mathcal{P}}^*$ and $F^*$ depend on the network node configuration, defined by $n_+$, $n_-$ and $n_p$ is now considered.

Each network must have at least one source node and one sink node, and is constrained by $n_+ + n_- + n_p = n$. These conditions restrict the space of possible network node configurations to a triangular-shaped simplex whose

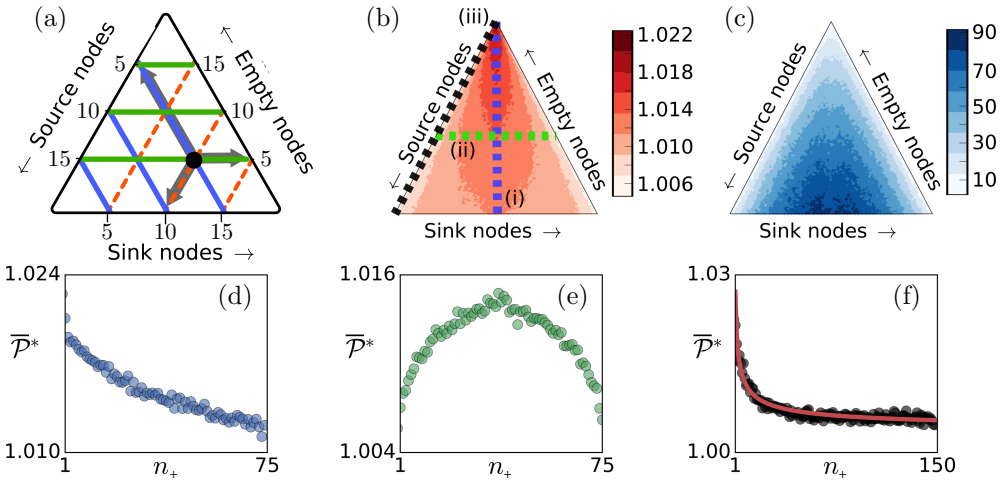



Figure 2.7: (a) A sketch of the node configuration space simplex. The black dot represents a configuration of $(n_+, n_-, n_p) = (5, 10, 5)$. (b), (c): $\overline{\mathcal{P}}^*$ and $F^*$, respectively, for an ensemble of 500 small world networks, each with $n = 150$, $k = 4$ and $q = 0.1$ projected onto the simplex in (a). (d)–(f) $\mathcal{P}$ as a function of $n_+$ along the sections (i)–(iii) indicated in (b). In (f), the red line indicates the function $a + bn_+^{-1/2}$ with $a = 1.003$ and $b = 0.024$.

vertices touch one of the $n_+, n_-, n_p$ axes, as illustrated in figure 2.7(a). Each unique node configuration $(n_+, n_-, n_p)$ is represented by a unique point on the simplex. For example the black dot in figure 2.7 represents a configuration of $(n_+, n_-, n_p) = (5, 10, 5)$. For networks with identical $n$, the values of $\overline{\mathcal{P}}^*$ and $F^*$ for each of the $n(n-1)/2$ possible configurations are then projected onto this simplex. The results of this projection for an ensemble of small-world networks with $n = 150$ and $q = 0.1$ are shown in figures 2.7(b) and 2.7(c). The contours in these results are symmetric about a line bisecting the simplex, corresponding to networks with equal numbers of sources and sinks; $n_+ = n_-$. This symmetry is due to the reciprocity between sources and sinks.

The maximum value of $\overline{\mathcal{P}}^*$ is found at the apex of the simplex in figure 2.7(b), corresponding to networks with only one source and one sink; i.e., $(n_+, n_-, n_p) = (1, 1, 148)$. The values of $\overline{\mathcal{P}}^*$ then decrease monotonically down the central bisecting line, indicated by section (i) in fig. 2.7(b) and plotted separately in fig. 2.7(d). Moving down this line, from the apex to

the bottom of the simplex, corresponds to increasing the numbers of sources and sinks in the network whilst keeping the proportions of them matched at $n_+ = n_-$. In other words, the networks become increasingly filled with an equal mixture sources and sinks. This decrease in $\overline{\mathcal{P}}^*$ with decreasing empty nodes is in agreement with the results found on lattice networks in figure 2.2; specifically, that efficiency is increased by adding in more sources and sinks. This increase in the efficiency of the Nash flow is due to the increase in source-sink nodes restricting the number of possible flow patterns, limiting the possible difference between the Nash and the GO.

Lines of constant numbers of passive nodes are represented as horizontal sections on the simplex, such as the section (ii). For each such line, the maximum $\overline{\mathcal{P}}^*$ is given in the centre where $n_+ = n_-$. Section (ii) is plotted separately in (e), and shows this to be the case; $\overline{\mathcal{P}}^*$ is maximised in the centre. This means that inefficiency is maximised when the numbers of sources and sinks are matched, for all values of passive nodes. Moving down section (iii) in fig. 2.7(b) corresponds to networks with only one sink, but increasingly many source nodes. As the source nodes are added in, $\overline{\mathcal{P}}^*$ decreases rapidly. This is shown in panel (f), revealing that $\overline{\mathcal{P}}^*$ decreases like $\sim a + bn_+^{-1/2}$.

In contrast to $\overline{\mathcal{P}}^*$, the value of $F^*$ is maximised at the bottom-centre of the simplex where there are no empty nodes. This is shown in fig. 2.7(c) and means that the fewer empty nodes there are, the higher the flow volume required to maximise inefficiency. Taking the results of fig. 2.7(b) and (c) together, it can be seen that although the highest inefficiencies occur when there is only one source and one sink, the flow volume at which this maximum occurs is very small. On the other hand, having a network half populated with sink nodes and its other half with sources, results in a small peak in inefficiency but occurring at a high flow volume. Therefore, the most efficient configuration of source-sink nodes depends upon the flow volume that the network will be handling.

The morphology of the contours shown in figures 2.7(b) and 2.7(c) remains invariant with the network rewiring parameter $q$, meaning these results apply to both small-world and random Poisson ($q \gtrapprox 0.2$) networks, as demonstrated in figure 2.8. These invariant properties also persist when considering scale-

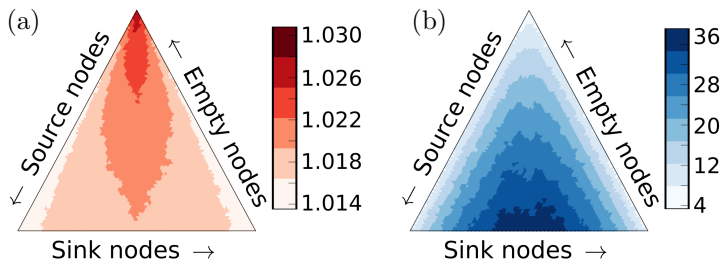



Figure 2.8: (a) $\overline{\mathcal{P}}^*$ and (b) $F^*$ for an ensemble of 500 random Poisson networks each with $n = 64$, generated by the Watts-Strogatz method (Watts and Strogatz, 1998) with $q = 0.6$ and $k = 4$.

free networks formed through preferential attachment, whose topologies are quite distinct from those of either the small-world or Poisson class. Figures 2.9(a) and 2.9(b) show the $\overline{\mathcal{P}}^*$ and $F^*$ simplexes for a scale-free network with $n = 150$ and mean degree 3, constructed using the Albert–Barabasi algorithm (Barabási and Albert, 1999). As in the case of small-world networks, the highest values of $\overline{\mathcal{P}}^*$ are found at the apex of the simplex and occur at a low flow volume; networks at the bottom of the simplex have lower maximum $\overline{\mathcal{P}}$ but occurring at a higher flow volume.

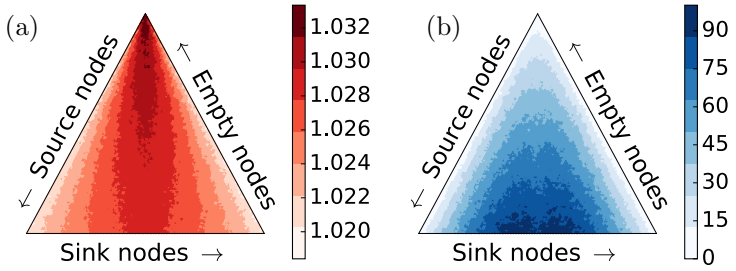



Figure 2.9: (a) $\overline{\mathcal{P}}^*$ and (b) $F^*$ for an ensemble of 500 scale-free networks with $n = 150$ and mean degree 3.

In practice sources and sinks may be expected to have heterogeneous levels of output and input, such as an electrical grid containing a range of generators with different output capacities. To account for this, $\xi_v$ in equation (2.4) is now set to be a normally distributed random variable with mean 0 and variance 0.2. This represents a substantial amount of heterogeneity whilst typically still preserving the types of the nodes, and therefore the location

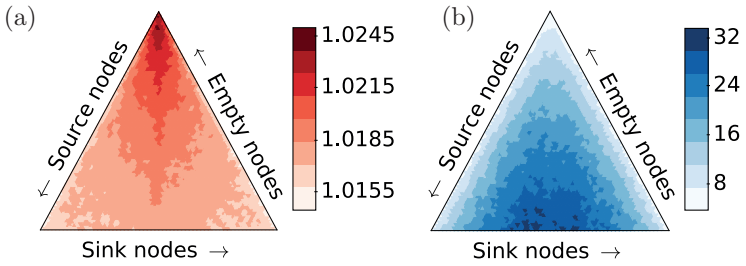



Figure 2.10: (a) $\overline{\mathcal{P}}^*$ and (b) $F^*$ for an ensemble of 500 random $n = 64$ small-world networks with $q = 0.1$. In these networks $\xi_v$ is a normally distributed random variable, with mean 0 and variance 0.2, inducing sources and sinks to have heterogeneous flow inputs and outputs.

on the simplex. Figure 2.10 demonstrates this heterogeneity in ensembles of small-world networks and reveals that the key features of the simplex remain. In particular the highest values of $\overline{\mathcal{P}}^*$ are found at the apex and along the centre line of the simplex where the numbers of source and sink nodes are equal.

The morphology of the contours in the simplex also remains approximately invariant with network size $n$, however, the amplitudes of $\overline{\mathcal{P}}^*$ and $F^*$ do scale with network size. To capture this scaling, a point in the centre of the simplex corresponding to networks with $n_+ = n_- = n_p$ was recorded for increasing $n$. These results are shown in figure 2.11. The value of $\overline{\mathcal{P}}^*$ as a function of $n$ is shown in fig. 2.11(a) and demonstrates that for small-world, Poisson and scale free networks it saturates to a constant value for $n > 50$. It is notable that $\overline{\mathcal{P}}^*$ saturates more quickly with $n$ for scale-free networks. This is likely because the characteristic hub-and-spoke structure of Barabási–Albert networks (Barabási and Albert, 1999) emerges at lower values of $n$ than does the small-world or Poisson quality of Watts-Strogatz networks. It is, however, unclear why $\overline{\mathcal{P}}^*$ saturates to a higher value for scale-free networks than the other networks. Figure 2.11(b) shows that $F^*$ increases linearly with network size. These scaling results can be used in conjunction with figure 2.7 to interrogate networks of arbitrary size.

The linear scaling shown in figure 2.11(b) can be explained. $F^*$ corresponds to a threshold beyond which the network flows adjust such that the two

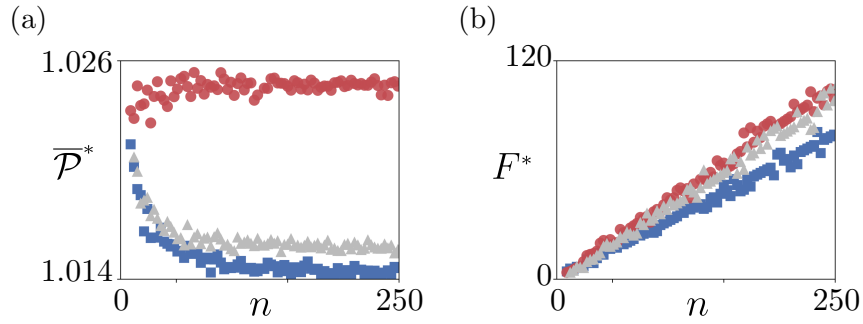



Figure 2.11: (a) $\overline{\mathcal{P}}^*$ and (b) $F^*$ as functions of $n$ for small-world networks (blue squares) with $q = 0.1$ and $k = 4$; Poisson networks (gray triangles) generated using the Watts–Strogatz method with $q = 0.6$ and $k = 4$ and scale-free networks (red circles), generated using the Barabási–Albert method (Barabási and Albert, 1999). All networks are chosen to have a node configuration $n_+ = n_- = n_p$.

equilibrium costs begin to converge. To exceed the threshold the total flow must increase linearly because the expected density of flow decreases linearly with increasing $n$.

## 2.5 Conclusion

This chapter has investigated how the inefficiency of flows occurring on different classes of random network, as gauged by the Price of Anarchy $\mathcal{P}$, is affected by the network structure and the function of its nodes. It has also established a correspondence between $\mathcal{P}$ and a measure of network redundancy, an important consideration in addressing issues of network resilience and cost-effectiveness. In particular, high values of $\mathcal{P}$ have been shown to indicate high levels of redundant capacity in a network.

This investigation has been primarily motivated by understanding properties associated with flows of current in electrical micro-grids, wherein nodes are either sources or sinks of current, or are passive conduits. Poisson, scale-free and small-world networks have been used to establish the generality of the results with respect to network topology; this reveals a predictable

dependence of $\mathcal{P}$ on node composition for networks of arbitrary structure.

The simplex plots of figure 2.7 and their symmetry and invariance properties, when taken in conjunction with the system size scalings shown in figure 2.11, provide an operating space that defines maximal inefficiency and redundancy for an ensemble of networks with general topology and with variable node composition. With application to micro-grids, a given network's composition will change both diurnally and seasonally, traversing a trajectory through this configuration space. This path will depend on the nature of the sources of power and the load consumed by the sinks. These features will vary with population behavior and the variable outputs from renewable power sources, as will be explored in Chapter 6. This information can be exploited to aid in the dynamic design and management of smart networks so as to constrain trajectories to preferred regions on the simplex. Insofar as redundancy is related to resilience (Halu et al., 2016; Quattrociocchi and Caldarelli, 2014; Corson, 2010), this aspect of the system's performance can be manipulated dynamically via the network's node type configuration and edge costing. A striking feature is that greatest values of inefficiency (or redundancy) occur when the number of sources and sinks are equal, as apparent in figure 2.7, a situation that is prevalent for small renewable energy networks where the numbers of generators and consumers are comparable. By contrast, the results show that a centralised electrical distribution grid comprising a few sources but many sinks has a low $\overline{\mathcal{P}}^*$, indicating it is both efficient and lacks redundancy. Equivalent plots can be constructed that are particular for an individual network's structure and composition with which its performance can be gauged. For instance, the redundancy in a micro-grid could be reduced by attempting to move the grid into a peripheral region of the simplex. Methods to accomplish this type of movement will be considered in Chapter 6.

These findings have established that even for simple linear edge functions, network topology and flow conservation laws are sufficient to induce inefficiency that depends predictably on the configuration and behaviour of nodes. An interesting extension to this work would be the consideration of nonlinear cost functions, for which the values of $\mathcal{P}$ may be substantially

larger (Roughgarden, 2003; Youn et al., 2008).

The inefficiency caused by redundancy is only one metric with which to assess performance and it is inefficient networks that will generally also be the most resilient to faults or attack. Redundancy may also give networks flexibility to operate in a variety of conditions; however, since inefficiency and redundancy coincide, optimising a network's structure and composition purely for efficiency may result in a loss of useful redundancy. Hence in using the simplex to aid network design it is likely that options will be constrained to an operating space offering an acceptable efficiency-resilience trade-off. The dependence of resilience upon node composition and behaviour forms the main theme of subsequent chapters.

# Chapter 3

# Resilience of steady state flow networks

## 3.1 Introduction

Cascading failures on networks can be catastrophic. A small fault or perturbation can spread throughout a system and compromise entirely its ability to operate. Blackouts in electrical grids, financial crashes and shocks, and the spread of congestion and delays through transport networks are all examples of cascades. Understanding how susceptible networks are to such events, and how certain topologies and network structures may improve resilience, is therefore vitally important to ensure the safe operation of infrastructure.

This chapter is concerned with cascades through steady state flow networks, which serve as a model for electrical power grids. Modern power grids have an increased number of small distributed generators, compared to traditional grids which contain only a few very large generators. Additionally, the composition and relative numbers of generators and consumers may vary throughout a day. Chapter 2 highlighted that inefficiency, gauged using the Price of Anarchy, varied as a function of the composition of flow source and sink nodes. The focus of this chapter is on how the resilience to cascades varies as a function of this source-sink composition in networks, in order

to uncover operating spaces that can boost the robustness of a network. The work presented here was published in Smith et al. (2020). Section 3.2 introduces the cascade model that will be used throughout the Chapter. Section 3.3 examines cascades on simple, regular lattice networks to develop general principles of how a cascade relates to edge capacity and node type. Finally, section 3.4 investigates cascades on a variety of complex networks, both synthetic and real, and computes their resilience as a function of their node-type composition.

## 3.2 Steady state cascade model

This section describes the network cascade model which will be used to investigate the resilience of flow networks to catastrophic failures. At every step of the cascade, the network flow is computed using the steady-state model introduced in section 1.5.1. As in Chapter 2, we consider networks of the form $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with $n = |\mathcal{V}|$ nodes and $m = |\mathcal{E}|$ edges, wherein $n_+$ nodes are sources of flow, $n_-$ are sinks and the remaining $n_p$ are passive or empty. Each edge $e \in \mathcal{E}$ has a flow volume $f_e$, computed using the linearised DC power flow equation (Stott et al., 2009)

$$f_e = \frac{\langle E_e, \theta \rangle}{x_e}. \tag{3.1}$$

Here $x_e$ is the reactance of edge $e$ and as before $E \in \mathbb{R}^{n \times m}$ is the node-edge incidence matrix, with $E_e$ its $e^{\text{th}}$ column. The inner product is denoted by $\langle \cdot, \cdot \rangle$ and $\theta \in \mathbb{R}^n$ is the vector of node voltages. The inner product in equation (3.1) determines the voltage difference across edge $e$, which is then divided by reactance to obtain a linearised current flow $f_e$. The node voltage vector $\theta$ is determined from the linear system

$$\mathcal{L}\,\theta = P, \tag{3.2}$$

where $\mathcal{L} = E\,\mathrm{diag}(Y)E^T$ is the weighted graph Laplacian and $Y \in \mathbb{R}^m$ the vector with entries $1/x_e$. $\mathcal{L}$ is also referred to as Kirchoff's matrix and finding a solution to equation (3.2) amounts to solving Kirchoff's voltage and

current laws. Note that these solutions are also equivalent to the Nash flow equilibria that were computed in chapter 2; however, those were computed using a convex optimisation algorithm, rather than the direct method applied here. The flow injection vector $P \in \mathbb{R}^n$ has entries

$$P_v = \begin{cases} (1+\xi_v)P^{\text{tot}}/n_+, & \text{if node } v \text{ is a source,} \\ -(1+\xi_v)P^{\text{tot}}/n_-, & \text{if node } v \text{ is a sink,} \\ 0, & \text{otherwise,} \end{cases} \tag{3.3}$$

with $P^{\text{tot}}$ the total power injected into the network and $\xi_v$ a random variable used to induce heterogeneity in source and sink strength. Supply and demand will always be matched so that

$$\sum_{v \in \mathcal{V}} P_v = 0. \tag{3.4}$$

It is worth noting that for every connected component in the network, the Laplacian will have a zero eigenvalue. If a network is simple and connected, then it will be composed of a single connected component and its Laplacian will therefore have rank $n-1$. This means that $\mathcal{L}$ in equation (3.2) is singular and non-invertible. However, computation of $f_e$ requires only the relative differences in voltage values between adjacent nodes. One of the node voltages may therefore be set to zero, and all other voltages measured relative to that ground node. This allows the system in (3.2) to be reduced to dimension $n-1$. This reduced system has full rank and can be solved directly via a LU decomposition.

Solutions of equations (3.1) and (3.2) describe a steady-state equilibrium flow which will be computed at each step of the network cascade, which proceeds as follows. First, a maximum capacity of $\alpha$ is assigned to each edge of the network. Next, the initial equilibrium flow is computed. The highest flow volume in the network at this stage will be denoted $\alpha_*$:

$$\alpha_* := \max_{e \in \mathcal{E}} f_e^{\text{initial}}. \tag{3.5}$$

$\alpha_*$ is the minimum edge capacity required for the network to function. Once

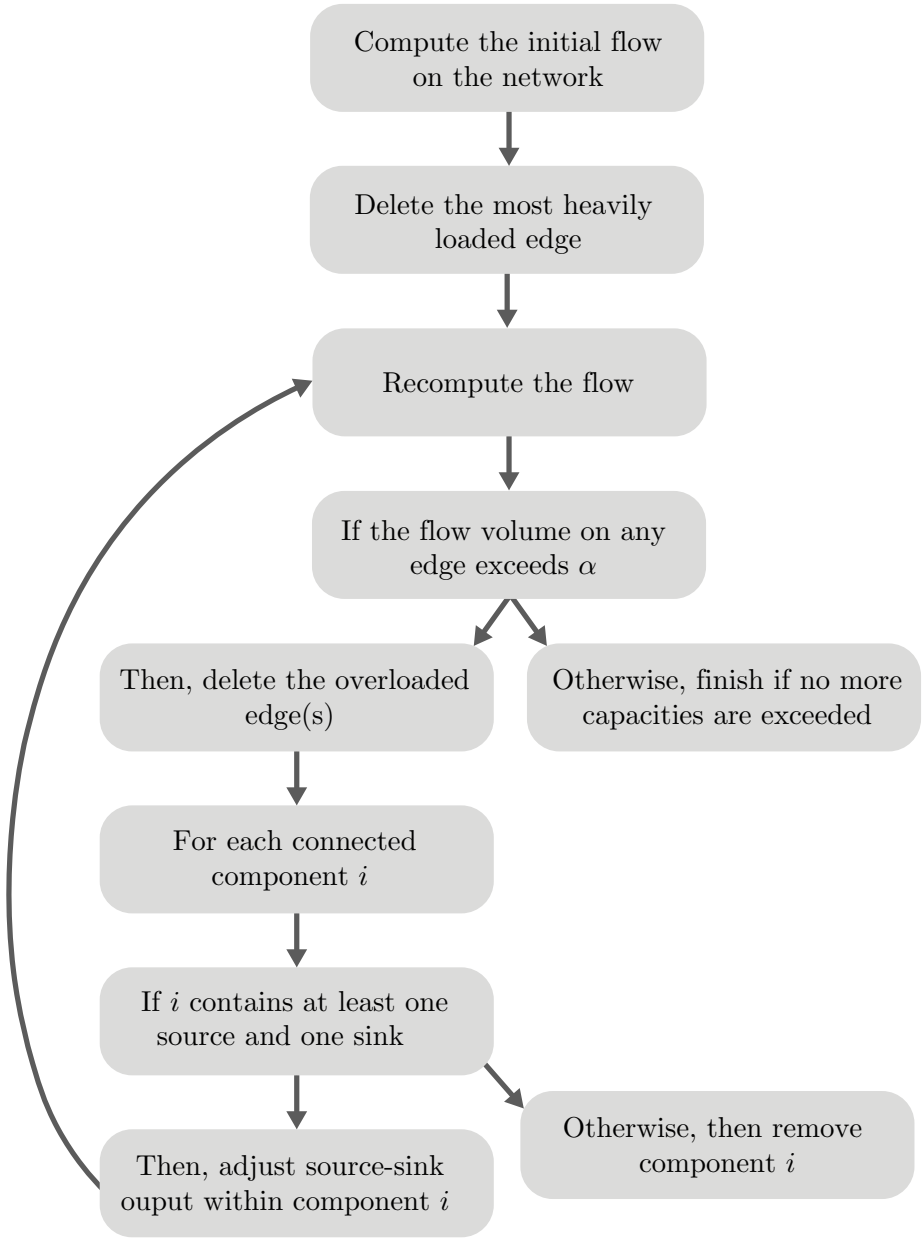



Figure 3.1: Schematic of the network cascade algorithm. The flow is computed by solving equations (3.1) and (3.2).

the initial flow is computed, the cascade is triggered by removing the edge with the highest flow volume. This serves as a model for overloading power line failure. The flow $f_e$ is then recalculated for each edge in this modified network. If $f_e > \alpha$ on any edge, that edge is deemed overloaded and removed. The flow is then recalculated and the network checked again for more overloads.

This process continues until the network attains a final equilibrium with all edges carrying a flow $\leq \alpha$.

During this cascade the network may fragment into separate connected components, and so the algorithm must be recursively forked into each of these new components. Additionally, source and sink strengths must be adjusted throughout the cascade in order to keep supply and demand matched, satisfying equation (3.4). It will be assumed initially that sources and sinks adjust their demand and supply reciprocally. In particular, if a component contains $n'_+$ sources, $n'_-$ sinks and a flow supply surplus $\delta$, then each source decreases its output by $\delta/2n'_+$ and each sink increases its demand by $\delta/2n'_-$. If the component contains either no sources or no sinks then the flow cannot be balanced and the entire component is removed. Balancing of power by non-reciprocal source/sink behaviour is discussed in section 3.5.1, together with other models for triggering the cascade. The cascade process is summarised in figure 3.1.

The network eventually attains a final state, with the fraction of surviving edges relative to the original number denoted by $\mathcal{S}$. A value of $\mathcal{S} = 0$ indicates an entire network failure; $\mathcal{S} = 1$ denotes complete resilience. As the edge capacity $\alpha$ in the network increases, the network will be better equipped to survive a cascade. Each network will require some amount $\alpha_*$ of capacity in order to operate normally. Some networks, owing to their structure and composition, will be inherently well equipped to survive cascades. Others, however, will be more fragile and require a large amount of capacity in excess of $\alpha_*$ in order to survive. This prompts defining a resilience metric $\rho$ that gauges how much capacity is needed for at least half of the network to survive, denoted $\alpha_c$, relative to the network's minimum operating capacity $\alpha_*$:

$$\rho := \frac{\alpha_c}{\alpha_*} \equiv \frac{\mathcal{S}^{-1}(1/2)}{\max_{e\in\mathcal{E}} f_e}. \tag{3.6}$$

When a network has a capacity of $\alpha_c$, half of the network will survive and so $\mathcal{S} = 1/2$. The smaller the value of $\rho$, the more resilient the network. If $\rho = 1$, for example, then the minimum operating capacity $\alpha_*$ alone is sufficient for most of the network to survive, indicating that the network

is inherently resilient. If $\rho = 2$, the network requires double its operating capacity. $\rho$ will be used throughout this chapter to gauge the resilience of a variety of networks.

## 3.3 Cascades on regular lattices

### 3.3.1 Lattices with degree $K = 2$

The fraction of surviving edges $\mathcal{S}$ depends on network size, topology, edge capacity, and both the composition and location of the node types. To gain insight into this dependence it is instructive to consider the example of a cascade on a network with a simple ring-lattice structure, as illustrated in figure 3.2(a). This simple lattice network admits analytic descriptions of its behaviour at each iteration of the cascade. We will now proceed to derive these descriptions and use them to reveal the dependence of $\mathcal{S}$ on the placement of the source nodes and the edge capacity values $\alpha$. It will then be shown that these results provide intuition for the general case of lattices with $K > 2$. Section 4.4 will also reveal that this intuition carries over even for the case of complex networks.

As shown in figure 3.2(a), the lattice has $n_+ = 2$, $n_- = n - 2$ and common degree $K = 2$. The shortest path distance between the two source nodes is $d$. For simplicity, let $\xi_v = 0$ for all $v$ in equation (3.3) so that sources and sinks are homogeneous. The power injected into each of the 2 source nodes is therefore $P^{\text{tot}}/2$, and $P^{\text{tot}}/(n-2)$ is consumed by each of the sinks. Here and throughout this chapter, the electrical flow is determined using the steady state equations (3.1) and (3.2). As described in Section 3.2, cascades are triggered by removing the most heavily loaded edge and then iteratively recomputing the flow and removing any further edges where $\alpha$ is exceeded. Initially, the maximum flow volume on any edge is

$$\max_{e \in \mathcal{E}} f_e = \frac{P^{\text{tot}}}{2} \left( 1 - \frac{d}{n-2} \right), \tag{3.7}$$

indicated by the red edges in figure 3.2(a). This is also $\alpha_*$, the minimum edge capacity required for the network to operate under normal conditions. One of these most loaded edges is knocked out to initiate the cascade. After

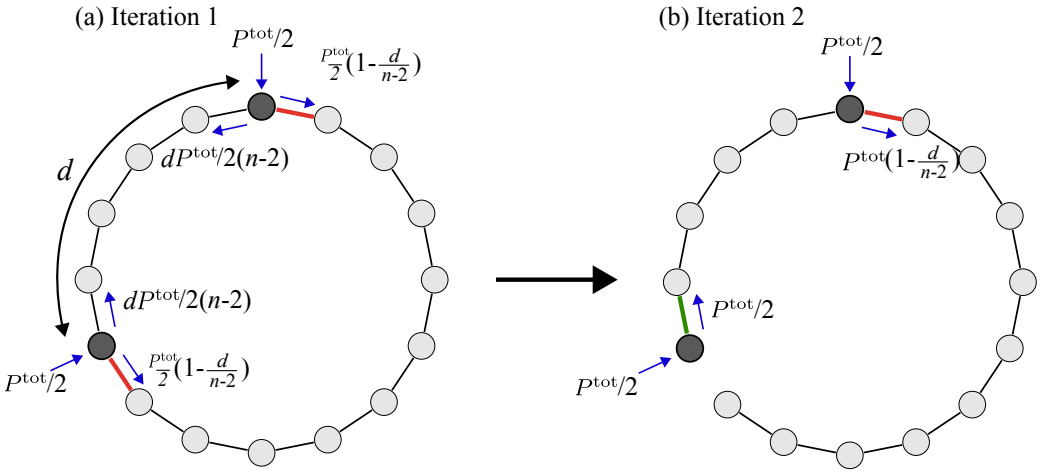



Figure 3.2: (a) Illustration of the ring-lattice with common degree $K = 2$, $n_+ = 2$ and $n_- = n - 2$. The distance between the two sources is labelled $d$. The annotations give the flow volumes on various edges of interest under the initial flow conditions. The two most loaded edges are highlighted red. (b) shows the second iteration of the cascade process, occurring after one of the most loaded edges in (a) is removed. After recomputing the flow, the new most loaded edge is highlighted in red.

this edge removal, the flow must then be carried by the one outgoing edge of the incident source node, as indicated by the green edge in figure 3.2(b). The most heavily loaded edge is now located where the flow from the first source mixes with the flow from the second, giving an edge flow volume of

$$f_2 = P^{\text{tot}} \left(1 - \frac{d}{n-2}\right), \tag{3.8}$$

indicated by the red edge in figure 3.2(b). If the edge capacity $\alpha \geq f_2$, then the edge can survive and the network can continue to operate with no further overloads. Thus the fraction of surviving edges is

$$\mathcal{S} = \frac{m-1}{m} \quad \text{for} \quad \alpha \geq f_2. \tag{3.9}$$

After the initial trigger, and subsequent redistribution of flow, some of the edges can still survive even if $\alpha < f_2$. In particular, the $d$ edges along the shortest path between the two sources survive, provided $\alpha$ exceeds the largest flow volume found along this path. This volume is $P^{\text{tot}}/2$, shown by the green edge in figure 3.2(b), and so

$$\mathcal{S} = \frac{d}{m} \quad \text{for} \quad \frac{P^{\text{tot}}}{2} \leq \alpha < f_2. \tag{3.10}$$

The resulting final state is illustrated in figure 3.3(b).

Whether any edges survive the cascade if $\alpha < P^{\text{tot}}/2$ depends upon the flow-balancing behaviour of the source and sink nodes; that is, if a connected component of the graph has a flow surplus or deficit, how do the source and sink nodes adjust their net flow outputs to match supply and demand? It is assumed for simplicity that this behaviour is reciprocal between sources and sinks. In particular, if a connected component has $n'_+$ sources and $n'_-$ sinks, and has a flow surplus of $\delta$, then each source will decrease its output by $\delta/2n'_+$ and each sink will increase its demand by $\delta/2n'_-$. In this way supply and demand will be equalised.

Under these assumptions, and if $\alpha < P^{\text{tot}}/2$, then a cascade will unfold as follows. First, the most loaded edge will be deleted. This will be an edge adjacent to one of the two sources, call this source 1, leading to the second iteration shown in figure 3.2(b), where the flow will be redistributed amongst the $m - 1$ remaining edges. The remaining edge incident on source 1 is then deleted (green edge in figure 3.2(b)), since its flow volume is now $P^{\text{tot}}/2 > \alpha$. The most loaded edge incident on source 2 with flow volume given by equation (3.8) will also be deleted (red edge in figure 3.2(b)), disconnecting most of the network. The remaining edge incident on source 2 lies along the shortest path between the two sources and will have a load given by

$$f_0 = P^{\text{tot}} \left( \frac{1}{2} - \frac{d}{n-2} \right). \tag{3.11}$$

If $\alpha \geq f_0$, then at least one edge can survive the second iteration of the cascade. As we move along this path, from source 2 towards source 1, the

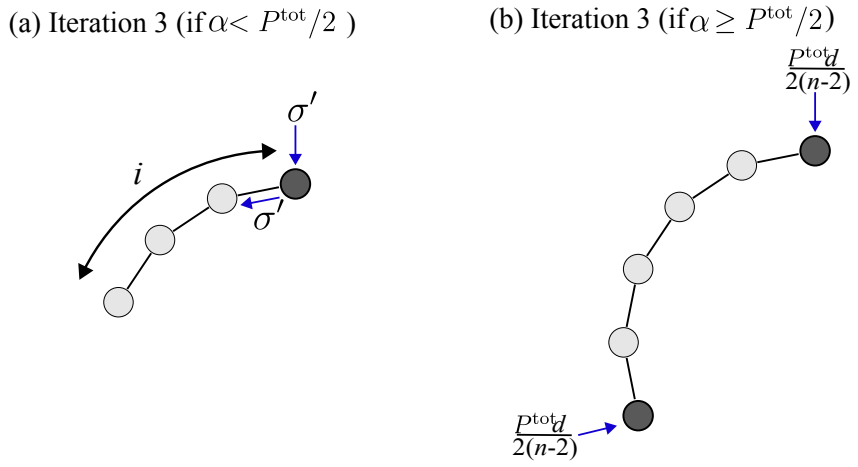



Figure 3.3: (a) and (b) show two possible outcomes of iteration 3 of the cascade, occurring after the red edge in figure 3.2(b) is removed due to $\alpha$ being less than $f_2$ in equation (3.10). (a) shows the outcome if $\alpha < P^{\text{tot}/2}$. The number $i$ of sink nodes surviving this far is given by the function $g(\alpha, d)$ in equation (3.13). These can survive the entire cascade if the relevant criteria in equation (3.18) are met. (b) shows the case if $\alpha \geq P^{\text{tot}}/2$.

$i$th edge along the path will have flow volume

$$f_i = P^{\text{tot}} \left( \frac{1}{2} - \frac{d-i}{n-2} \right). \tag{3.12}$$

A chain of $i$ edges as shown in figure 3.3(a) can therefore survive the second round of the algorithm if $\alpha \geq f_i$. Thus, for these values of $\alpha$, the number of edges $i$ surviving this round of the cascade is obtained by letting $f_i = \alpha$ in equation (3.12) and rearranging for $i$. After some manipulation this yields the result

$$i = r - (n-2) \left( \frac{1}{2} - \frac{\alpha}{P^{\text{tot}}} \right) =: g(\alpha, d). \tag{3.13}$$

The fraction of edges surviving this iteration is then $\lfloor g(\alpha, d) \rfloor / m$, where $\lfloor \cdot \rfloor$ denotes the floor function. Now, assuming at least one edge has survived so far, the third iteration of the algorithm will begin, which will start by balancing the supply and demand on the remaining network. This remaining structure will consist of one source, with output $\sigma = P^{\text{tot}}/2$, and $i$ sinks, connected in a chain, each with demand $P^{\text{tot}}/(n-2)$. There will therefore

be a flow surplus of

$$\delta = \frac{P^{\text{tot}}}{2} - \frac{iP^{\text{tot}}}{n-2}, \tag{3.14}$$

and the source will decrease its output by $\delta/2$, giving a new output of

$$\begin{aligned}\sigma' &= \frac{P^{\text{tot}}}{2} - \frac{\delta}{2} = \frac{P^{\text{tot}}}{2} - \frac{1}{2}\left(\frac{P^{\text{tot}}}{2} - \frac{iP^{\text{tot}}}{n-2}\right) \\ &= \frac{P^{\text{tot}}(n-2) + 2iP^{\text{tot}}}{4(n-2)}. \end{aligned} \tag{3.15}$$

After balancing the supply and demand, the new flow pattern will be computed. The most loaded edge will now be the one incident on the source, and will have to carry all of the flow from the source to the other edges. This edge will therefore have a load equal to $\sigma'$ given in equation (3.15). If $\alpha$ is greater than or equal to this load $\sigma'$, then the $i$ edges in the chain survive and the network will have reached its final state with $\mathcal{S} = \lfloor g(\alpha, d) \rfloor / m$. Otherwise if $\alpha < \sigma'$, then $\mathcal{S} = 0$. A lower bound on $\alpha$ therefore occurs when

$$\alpha = \frac{P^{\text{tot}}(n-2) + 2iP^{\text{tot}}}{4(n-2)},$$

which upon rearranging for $i$ gives

$$i = \frac{(4\alpha - P^{\text{tot}})(n-2)}{2P^{\text{tot}}}, \tag{3.16}$$

where $i$ is given by $g(\alpha, d)$ in equation (3.13). Therefore, equating (3.13) and (3.16)

$$d - (n-2)\left(\frac{1}{2} - \frac{\alpha}{P^{\text{tot}}}\right) = \frac{(4\alpha - P^{\text{tot}})(n-2)}{2P^{\text{tot}}},$$

and rearranging and simplifying gives the result

$$\alpha = \frac{P^{\text{tot}} d}{(n-2)}, \tag{3.17}$$

which is the capacity required for the $\lfloor g(\alpha, d) \rfloor$ edges which survived the second iteration to also survive the third and final iteration. This final state

is illustrated in figure 3.3(a). Recall that for anything to have survived the second iteration, $\alpha$ must be greater than or equal to $f_0$ in equation (3.11). The results in equations (3.9), (3.10) and (3.17) can now be summarised as

$$\mathcal{S} = \begin{cases} 0 & \text{for} \quad (\alpha, d) \quad \text{where} \quad \alpha < P^{\text{tot}}\left(\frac{1}{2} - \frac{d}{n-2}\right) \text{ and } d > \frac{\alpha(n-2)}{P^{\text{tot}}}, \\ \frac{1}{m}\lfloor g(\alpha, d)\rfloor & \text{for} \quad (\alpha, d) \quad \text{where} \quad P^{\text{tot}}\left(\frac{1}{2} - \frac{d}{n-2}\right) \leq \alpha \leq \frac{P^{\text{tot}}}{2} \text{ and } d \leq \frac{\alpha(n-2)}{P^{\text{tot}}} \\ \frac{d}{m} & \text{for} \quad (\alpha, d) \quad \text{where} \quad \frac{P^{\text{tot}}}{2} \leq \alpha < P^{\text{tot}}\left(1 - \frac{d}{n-2}\right) \\ \frac{m-1}{m} & \text{for} \quad (\alpha, d) \quad \text{where} \quad \alpha \geq P^{\text{tot}}\left(1 - \frac{d}{n-2}\right). \end{cases} \tag{3.18}$$

This provides a full description of cascades for a given capacity $\alpha$ and source distance $d$. The numerical simulation results in figure 3.4(a), showing $\mathcal{S}$ as a function of $d$ and $\alpha$, verify the results in equation (3.18). Figure 3.4(b) shows the same results, but where the horizontal axis has been normalised by $\alpha_*$. Recall from equation (3.5) that $\alpha_*$ is the minimum capacity required for

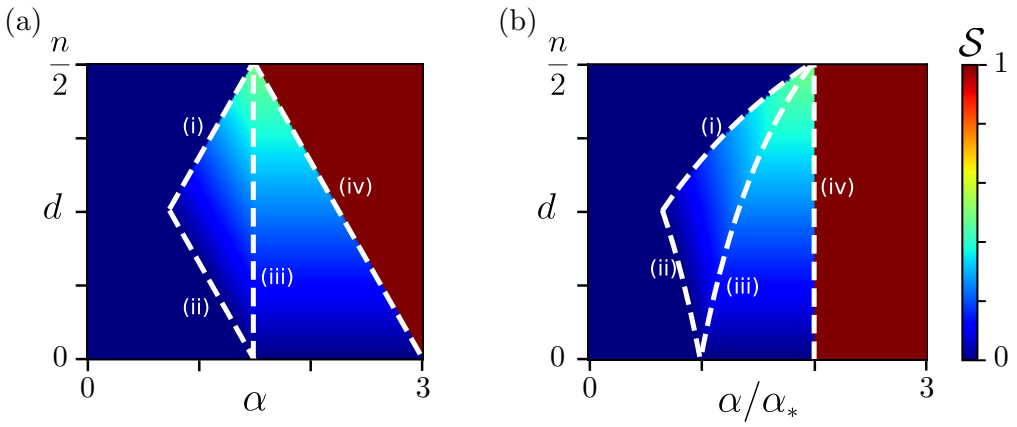



Figure 3.4: Panel (a) shows the fraction $\mathcal{S}$ of edges surviving a cascade on a ring-lattice network with $n$ nodes and with two source nodes separated by a distance $d$, as illustrated in figure 3.2(a). $\mathcal{S}$ is plotted as a function of $d$ and edge capacity $\alpha$. The white dashed boundaries labelled (i) to (iv), indicate sequentially the bounds in equation (3.18). Red indicates network survival. Panel (b) shows the same results as (a) but with the horizontal axis normalised by $\alpha_*$. The resilience boundary (iv) now occurs at a value of 2 for all $d$.

the network to be able to operate under normal conditions, and is equal to the maximum flow volume before the cascade occurs. The quantity $\alpha/\alpha_*$ is

thus the capacity in the network relative to the minimum required for normal function.

The normalisation in figure 3.4(b) reveals that for $\alpha/\alpha_* \geq 2$ the network is resilient to cascade failures, independent of the relative locations of the source nodes. This can be understood as follows. $\alpha_*$ is defined as the largest initial flow volume in the network, in this case equal to the quantity in equation (3.7). By comparing equations (3.10) and (3.7), it can be seen that $2\alpha_* = f_2$. The bound in equation (3.9) shows that all but one of the edges survive a cascade if $\alpha \geq f_2$. Normalising this by $\alpha_*$ therefore gives the required resilience boundary. For $\alpha/\alpha_* < 2$, the network is less robust with resilience conditional on $d$.

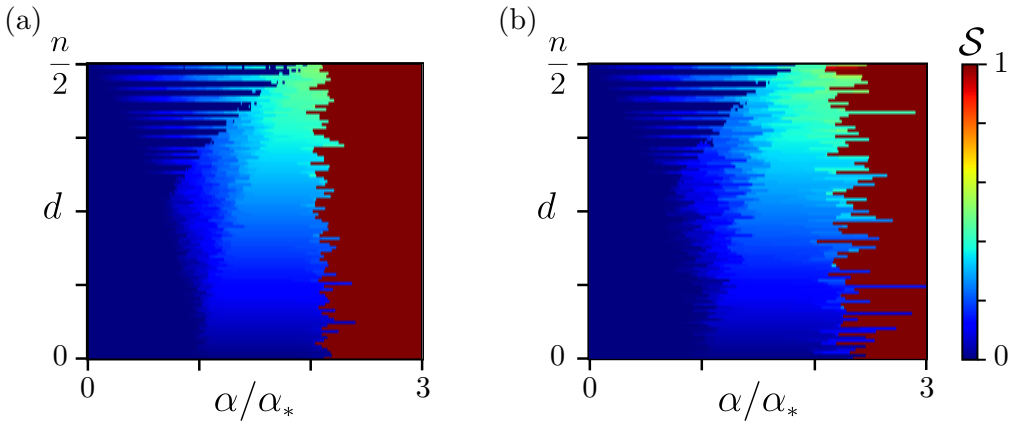



Figure 3.5: Fraction $\mathcal{S}$ of edges surviving a cascade in ring lattice networks, as illustrated in figure 3.2. Here the edge capacities have been perturbed by Gaussian noise with standard deviation 0.05 in (a) and 0.1 in (b). $\mathcal{S}$ is shown as a function of the distance $d$ between the two sources and the mean normalised capacity. Each row of the surfaces is a particular realisation of the network. The distinctive boundaries located in figure 3.4 are still visible.

It is useful here to employ the resilience metric $\rho = \alpha_c/\alpha_*$ defined in equation (3.6), where $\alpha_c$ is the capacity required for at least half of the edges to survive. In this case, we have $\rho = 2$ for all values of $d$. In other words, 2 times the minimum operating capacity is required for most of the network to survive.

To investigate the effect of stochasticity in the cascade process, figures 3.5(a)

and (b) show the case where the capacities $\alpha$ on each edge are perturbed with Gaussian noise. The features and boundaries identified in equation (3.18) approximately remain.

### 3.3.2 Lattices with degree $K > 2$

While lattices with $K = 2$ admit a simple analytic description of their full behaviour, lattices with $K > 2$ are more complex due to the presence of overlapping paths. We shall therefore rely on approximations and numerical simulation to determine their resilience to cascade failures.

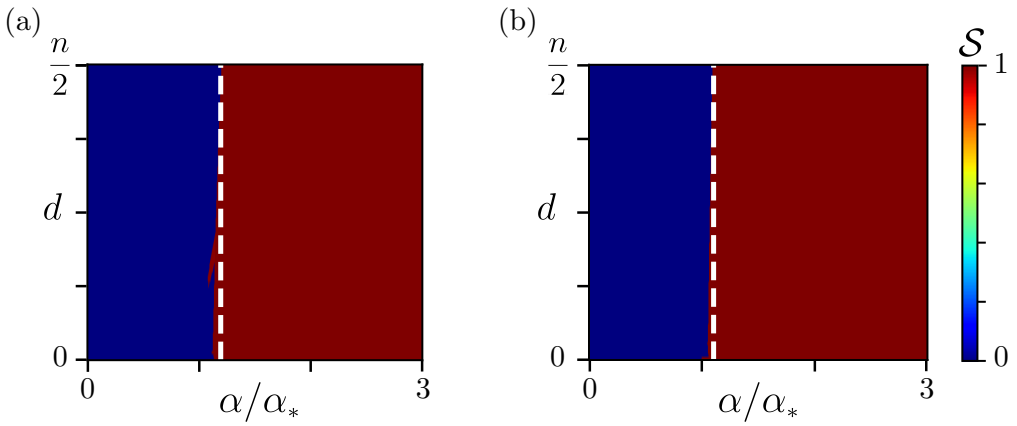



Figure 3.6: Panels (a) and (b) the fraction of edges $\mathcal{S}$ surviving a cascade on a ring lattice with degree $K = 6$ and $K = 10$ respectively. The white line indicates the resilience boundary, beyond which the network survives the cascade.

As $K$ increases, the conditional boundaries identified in equation (3.18) erode due to the presence of multiple paths along which the flow can equilibrate. Nevertheless, the resilience boundary, as seen in figure 3.4, remains robust and independent of source separation $d$. This is demonstrated in figures 3.6(a) and (b), which show the fraction of surviving edges $\mathcal{S}$ for lattices with $K = 6$ and $K = 10$ respectively. Beyond the boundary shown in these figures, most of the network survives. The location of the boundary with respect to $\alpha/\alpha_*$ therefore gives the value of $\rho$, the relative increase in capacity required for

most of the network to survive. As is evident from figures 3.6(a) and (b), the value of $\rho$ decreases as $K$ goes from 6 to 10. A decrease in $\rho$ indicates increased resilience, since less capacity is required for the network to survive.

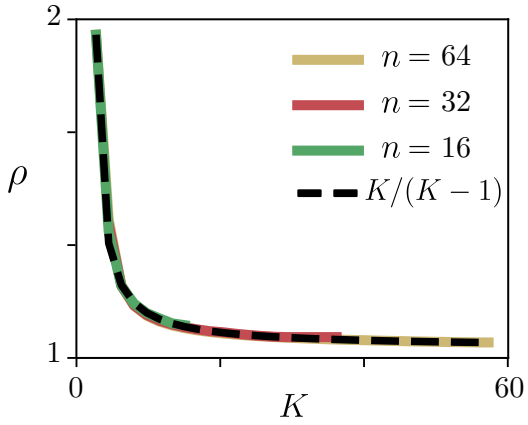



Figure 3.7: The resilience $\rho$, equivalent to the location of the resilience boundary, as a function of degree $K$ for a variety of lattice sizes $n$. The data is fitted with $\rho = K/(K-1)$.

To track how $\rho$ decreases, it is plotted as a function of $K$ in figure 3.7. The values of $\rho$ were computed numerically from simulations on lattices of different sizes $n$. This reveals that $\rho$ decreases approximately as $\rho \approx K/(K-1)$. The value of $\rho$ in lattices therefore rapidly decreases with $K$, meaning that increasing the degree of the lattice dramatically increases resilience.

This $K/(K-1)$ behaviour can be understood by considering the ring-lattice network with common degree $K$ when $n_+ = 1$ and $n_- = n - 1$. The initial flow in this case will be apportioned equally across the $K$ edges emanating from the single source. These edges are the most heavily loaded and so $\alpha_* \approx P^{\text{tot}}/K$. When one of these edges is removed to initiate a cascade, the flow is redistributed among the $K-1$ remaining edges connected to the source and these must now supply flow to the rest of the network. For the lattice to survive the cascade, there must be sufficient capacity in the system for these edges to survive. Consequently $\alpha_c \approx P^{\text{tot}}/(K-1)$ and therefore $\alpha_c/\alpha_* \equiv \rho \approx K/(K-1)$.

## 3.4 Cascades on complex networks

Having investigated the effects of cascade failures on lattices in the previous section, this section introduces some exemplar complex networks and describes how the resilience metric $\rho$ is calculated for such networks.

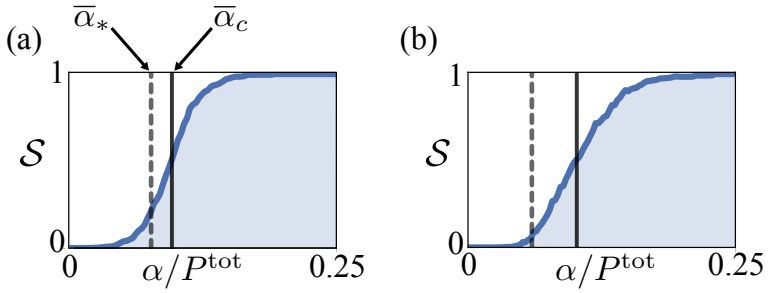



Figure 3.8: The mean fraction of edges $\bar{\mathcal{S}}$ surviving a cascade as a function of edge capacity $\alpha$ for ensembles of 200 Watts-Strogatz networks with $n = 50$, $K = 4$ and $(n_+, n_-, n_p) = (10, 40, 0)$. Source-sink node locations are chosen randomly. (a) and (b) are for $q = 0.1$ and $q = 1$ respectively. The mean minimum operating capacity $\overline{\alpha}_*$ and the mean critical capacity $\overline{\alpha}_c$, beyond which more than half of the network structure survives, are indicated by the dashed and solid lines respectively.

As before, the fraction of surviving edges $\mathcal{S}$ will be computed as a function of edge capacity using the algorithm outlined in figure (3.1), this time for ensembles of randomly generated complex networks. Figures 3.8(a) and (b) show the mean number of surviving edges $\bar{S}$ against $\alpha$ for two ensembles of random networks generated using the Watts–Strogatz method (Watts and Strogatz, 1998). The networks all have size $n = 50$, $m = 100$, and with source and sink node locations chosen uniformly at random. The chosen rewiring parameters are $q = 0.1$ in figure 3.8(a) and $q = 1$ in (b), characteristic of small-world and Poisson networks respectively. Both network ensembles are susceptible to substantial disruption for capacities $\alpha \lessapprox 0.06$. For $\alpha \gtrapprox 0.2$, they are both essentially robust with nearly all edges surviving. The ensemble mean operating capacity $\bar{\alpha}_*$ is marked in both (a) and (b), along with the mean critical capacity $\bar{\alpha}_c$ which is the capacity value beyond which more than half of the network will on average survive. Figures 3.8(a) and (b) show that $\bar{\alpha}_c$ increases as the networks become more Poissonian,

meaning that greater capacity is required for the networks to survive as they become more random. A bisection method is used to locate $\bar{\alpha_c}$ and $\bar{\alpha_*}$ from data of the type shown in figure 3.8, allowing the mean resilience metric $\bar{\rho}$ to be calculated. The relative location of $\bar{\rho}$, for different network structures and node compositions will be the focus of Section 3.5.

For ensembles of random networks $\rho$ does not take a single value but rather is a random variable. For Watts-Strogatz networks, we find that the probability density $P(\rho)$ is found to be log-normal:

$$P(\rho) = \frac{1}{\rho\nu\sqrt{2\pi}} \exp\left(-\frac{(\ln\rho - \mu)^2}{2\nu^2}\right), \tag{3.19}$$

with mean value $\mu = \bar{\rho}$ that scales with network size $n$ but with similar standard deviation $\nu \approx 0.255$ across the full range of rewiring parameters $q \in [0, 1]$, encompassing small-world to Poisson network topologies. This data collapse is demonstrated in figure 3.9(b) which shows the distribution of $\rho$ for ensembles of networks with $q = 0.1$ and $q = 0.6$, labelled (i) and (ii) respectively. The reason for this log-normality is possibly related to fragmentation processes; it is known that the sizes of fragments of an object after being repeatedly smashed apart are log-normally distributed (Cheng and Redner, 1988). For instance, sediment particle sizes after continued erosion (Dacey and Krumbein, 1979), and even the sizes of pieces of food after chewing (Kobayashi et al., 2006), have been identified as log-normal. This suggests the possibility that network cascades could be regarded as a type of fragmentation process. However, our objective here is only to fit a distribution in order for the data to be conveniently expressed using two parameters; this will allow for an efficient method of sampling later on in Chapter 6. Therefore, the relationship between cascades, fragmentation, and log-normal distributions will be not be investigated further in this thesis. Let us note, however, that it could form an interesting topic for future research.

Plots (iii) and (iv) in figure 3.9(b) show the probability densities for two ensembles of scale-free networks that have power-law degree distribution with initial node degree $m_0 = 3$ and $m_0 = 4$ respectively. These networks are generated using the Albert–Barabasi preferential attachment method

(Barabási and Albert, 1999). The range of $\rho$ is smaller than that found for Watts–Strogatz networks, thereby indicating a greater resilience, and this agrees with earlier studies on the resilience of scale-free networks predicated on random edge removal (Callaway et al., 2000; Albert et al., 2000). When $n_+ \approx n_-$ the distribution is bimodal, with the larger component of the distribution corresponding to when the edge triggering the cascade is connected to a hub. It may be possible that more peaks emerge in the distribution for larger scale-free networks, as more hubs with varying vulnerabilities begin to form in the network. Networks grown by preferential attachment, such as social networks and the Internet, possess a scale-free structure, whereas electrical grids and road/rail transportation networks do not because their structures are often dictated by geographical constraints. This chapter is concerned chiefly with electrical grids, so from here networks generated by the Watts–Strogatz procedure shall provide the structural substrate on which general results regarding resilience to failures can be deduced. The form of the resilience distributions for larger scale-free networks shall therefore be left for future research.

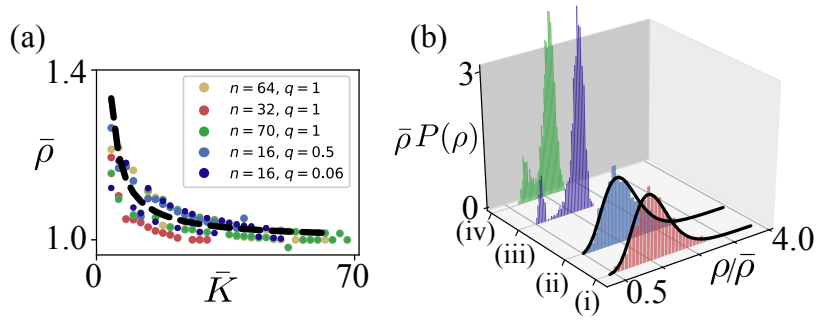



Figure 3.9: (a) shows the mean resilience $\bar{\rho}$ plotted against mean degree $\bar{K}$ for various ensembles of Watts–Strogatz networks, parameterised by their size $n$ and topological randomness $q$. The dashed black line is $\bar{\rho} = \bar{K}/(\bar{K}-1)$. (b) shows the probability density $P(\rho)$ in random networks with $(n_+, n_-, n_p) = (30, 30, 0)$. Histograms (i) and (ii) are for ensembles of 200 Watts–Strogatz networks with $q = 0.1$ and $q = 0.6$. Both are fitted to Lognormal$(\mu, \nu^2)$, with $(\mu, \nu) = (1.18, 0.25)$ in (i) and $(\mu, \nu) = (1.63, 0.26)$ in (ii). Both fits are confirmed with Kolmogorov–Smirnov tests Frank and Massey (1951). Histograms (iii) and (iv) are for scale-free networks with $m_0 = 3$ and $m_0 = 4$ respectively.

The approximation of $\rho \approx K/(K-1)$ uncovered in Section 3.3.2 can also

be detected in the case of random networks containing only one or two large source nodes, and whose remaining nodes are sinks. Figure 3.9(a) shows $\bar{\rho}$ plotted against $\bar{K}$ for a selection of such networks whose topologies are of Watts-Strogatz type. The data is overlaid with $\bar{\rho} = \bar{K}/(\bar{K} - 1)$ where $\bar{K}$ is the mean degree. $\bar{\rho}$ therefore decreases rapidly with $\bar{K}$ indicating that increasing degree drastically improves resilience for small-world and Poisson networks, as well as in lattices as seen in the previous section.

## 3.5 Dependence on node composition and function

This section investigates how the resilience of complex networks depends on the proportions of source and sink nodes. This is motivated by the rapid increase in the volume of renewable, distributed power generation in modern electrical grids. Understanding the role of source-sink composition is therefore vital to understand how the shift from centralised generation to smaller, more distributed generation impacts network function.

### 3.5.1 Synthetic complex networks

The resilience metric $\rho$ as defined in equation (3.6) will be computed as a function of the proportions of sources and sinks in a variety of complex networks. Following the approach of Chapter 2, $\rho$ will then be projected onto the node configuration simplex as illustrated in figure 3.10(a). Recall that the condition $n_+ + n_- + n_p = 0$, together with the requirement that there be at least one source and one sink node, constrains the space of possible node-type compositions onto this triangular simplex. The length of an edge of the simplex is the size $n$ of the network and each of the $n(n-1)/2$ unique points on the triangle represents a unique node-type composition. For example, the dot in figure 3.10(a) represents $(n_+, n_-, n_p) = (5, 10, 5)$. Ensembles of Watts–Strogatz networks, paramaterised by $(q, K, n)$, are used and node-types according to $(n_+, n_-, n_p)$ are placed uniformly at random. The mean resilience $\bar{\rho}$ is calculated from 200 realisations of such networks

and its value projected onto the corresponding location on the simplex. $\bar{\rho}$ is the mean increase in capacity required for more than half a network to survive a cascade; the lower the value of $\bar{\rho}$, the more resilient the network.

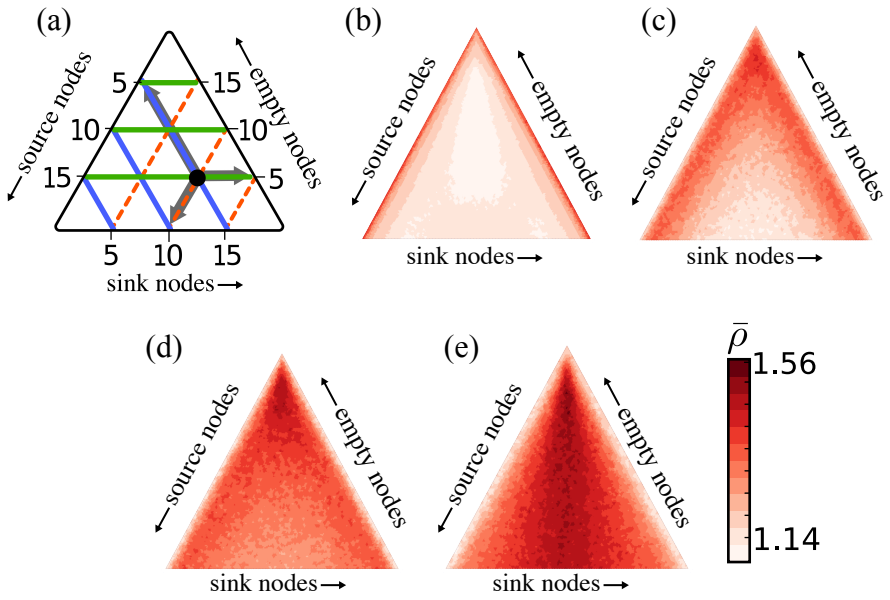



Figure 3.10: (a) A sketch of the node configuration simplex. The black dot represent a node-type configuration of $(n_+, n_-, n_p) = (5, 10, 5)$. Plots (b) to (e) show $\bar{\rho}$ projected onto the simplex for ensembles of 200 Watts-Strogatz networks of size $n = 100$, $K = 4$ and with $q = 0, 0.2, 0.4$ and 1 respectively. (b) to (e) therefore represent increasing topological randomness, from regular lattices to Poisson networks.

Figure 3.10(b) shows $\bar{\rho}$ for regular lattices ($q = 0$) with $n = 100$, $K = 4$, as a function of node-type composition The largest values of $\bar{\rho}$ and therefore the lowest resilience are located down two edges of the simplex, corresponding to networks with large numbers of sinks and very few sources, or vice versa. This means that the few source (or sink) nodes must generate (or absorb) the flow, and so incident edges must carry large flow volumes, making the network susceptible to failure. Resilient networks are spread throughout the interior region, corresponding to networks with many small, highly distributed sources and sinks. This agrees with the intuition that a mix of node-types has a high degree of redundancy and is therefore resilient because of the multiple flow paths such networks contain.

As $q$ increases and the network structure becomes more random the morphology of $\bar{\rho}$ changes, the trend being shown in figures 3.10(c) to (e). A band of greatest $\bar{\rho}$ broadens from the edges, moving into the interior region, until for $q \gtrapprox 0.12$ the largest values of $\bar{\rho}$ are found in the interior of the simplex. This represents an inversion in the resilience measure, counter-intuitively showing that for networks in this regime of $q$ a more heterogeneous mix of node-types give less resilience than does a network with a few large suppliers of power. This behaviour of $\bar{\rho}$ continues progressively until $q = 1$ (a Poisson network), where $\bar{\rho}$ has largest values in the centre of the simplex and lowest values along two edges. In all cases the morphology of $\bar{\rho}$ is symmetric about the centre line, reflecting the reciprocity between sources and sinks.

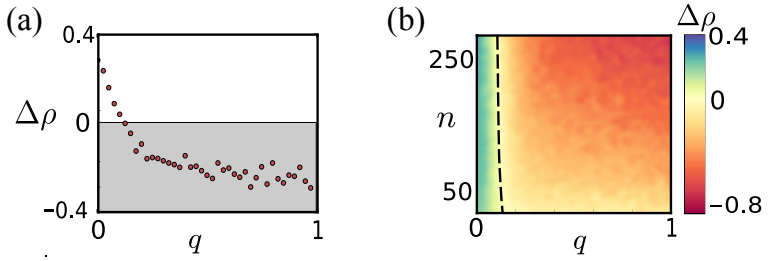



Figure 3.11: (a) $\Delta\rho$ as defined in equation (3.20) as a function of $q$ for ensembles of Watts–Strogatz networks with $n = 60$. (b) $\Delta\rho$ as a function of both $q$ and $n$. The dashed line is the locus of the inversion point $\Delta\rho = 0$. All data points are averaged over 200 realisations.

The inversion in the behaviour of $\bar{\rho}$ with $q$ can be gauged by determining the value of

$$\Delta\rho := \bar{\rho}_{\text{Left}} - \bar{\rho}_{\text{Centre}}, \tag{3.20}$$

where $\bar{\rho}_{\text{Left}}$ and $\bar{\rho}_{\text{Centre}}$ are the values of $\bar{\rho}$ at the bottom-left and centre of the simplex respectively. If $\Delta\rho > 0$, then the lowest values of $\bar{\rho}$ are to be found in the centre of the simplex, meaning that having many small distributed sources boosts resilience. If $\Delta\rho < 0$ then the greatest resilience, and so the smallest values of $\bar{\rho}$, are at the exterior of the simplex, corresponding to networks with only a small number of very large sources. If $\Delta\rho \approx 0$, then the resilience is approximately invariant to node-type composition.

For regular lattices $\Delta\rho > 0$, whereas for Poisson networks $\Delta\rho < 0$. Figure 3.11(a) shows that for networks with $n = 60$, $\Delta\rho$ decreases monotonically as

a function of $q$, the inversion occurring at the value of $q$ where the network transitions from approximate small-world order to the more Poissonian regime. Figure 3.11(b) shows the dependence of $\Delta\rho$ with $q$ and $n$. $\Delta\rho = 0$ is shown by the dashed line and denotes the locus of the inversion, which is approximately independent of network size. Therefore, for networks of any size, the greatest resilience is given by having a mix of many small sources and sinks when the networks are lattices or small-world. The opposite is true for more random Poisson networks; the greatest resilience results from having only a few large sources.

Preceding results consider homogeneous flow injection (*i.e.* $\xi_v = 0$ in equation (3.3)) and cascades triggered by the removal of the most heavily loaded edge, which serve as a model for overloading failure. The impact of varying these model choices, and thus introducing several different types of heterogeneity, will now be considered.

It is assumed so far that both sources and sinks have the same ability to adjust their outputs and inputs respectively to match supply and demand of flow as the network begins to break down. In practice, in electrical grids, different types of nodes may have differing abilities to adjust their operation. For example, when a cascading blackout begins, the grid will typically begin to shed loads, meaning sink nodes decrease their inputs to prevent further network damage. In other cases, if there is excess power on the grid, batteries may be brought online to increase demand. Both cases represent situations where the sinks (loads) have greater flexibility to adjust their demands in the timescales required than the sources (generators).

The case where only sink nodes have the ability to adjust their flow is considered in figure 3.12. The main features seen in the case of reciprocal balancing in figure 3.10 remain the same; the lowest values of $\bar{\rho}$ are are found in the centre of the simplex for regular networks and around the edge for more disordered ones. Indeed any such heterogeneity in the abilities of the nodes causes a break in the symmetry, where the lower values of $\bar{\rho}$ skew towards node configurations containing more of whichever type of node possesses the greatest adaptability. In the case of figure 3.12, the lowest values skew towards the side representing networks with high numbers of sinks, since

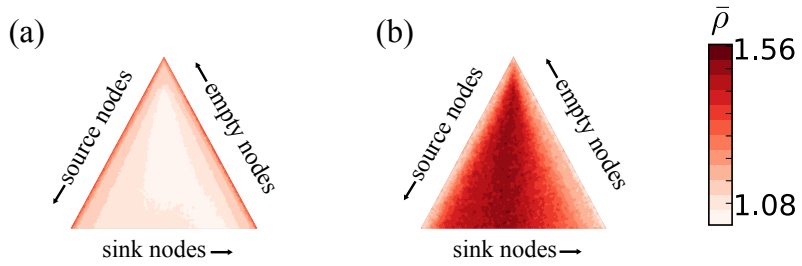



Figure 3.12: $\bar{\rho}$ for Watts–Strogatz networks with $n = 100$, $k = 4$ and $q = 0$ in (a) and $q = 1$ in (b). Each pixel represents an average over 200 network realisations. In both (a) and (b), only the sink nodes have the ability to adjust their flow in order to balance supply and demand. This results in a break in the symmetry of the simplex.

these are the node type with the greatest flexibility in this case.

A further assumption made so far is that cascades are triggered by the removal of the most heavily loaded edge. This serves as a model of overloading line failure, where more heavily loaded lines are more likely to suffer failure. There may, however, be other initial triggers such as lightning strikes or software failures. To take this into account, cascades are now modelled with the triggering edge chosen with a probability proportional to its initial load. In particular, if an edge initially has load $f_e$, then it is chosen as the trigger with a probability $|f_e|/||f||_1$, where $f \in \mathbb{R}^m$ is the vector of edge loads and $||\cdot||_1$ is the $L^1$ vector norm. Figures 3.13(a) and (b) show this case for Watts–Strogatz networks with $q = 0$ and $q = 1$ respectively. Again, the key features of the homogeneous case remain, including the inversion with respect to $q$.

So far it has also been assumed that source and sink output strengths are homogeneous; however, in reality there is often significant variation in strength. To account for this heterogeneity in source-sink size, $\xi_v$ in equation (3.3) is now set to be a normally distributed random variable and the sources and sinks now balance flow proportionally to their initial load. In particular if a network fragment contains $n'_+$ sources and $n'_-$ sinks and has flow surplus

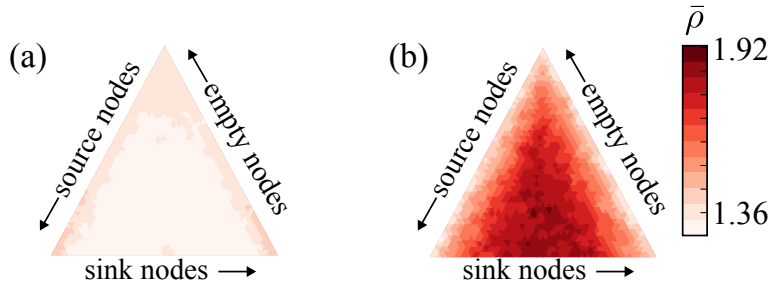



Figure 3.13: $\bar{\rho}$ for Watts–Strogatz networks with $n = 100$, $k = 4$ and $q = 0$ in (a) and $q = 1$ in (b). Each pixel represents an average over 200 network realisations. In both (a) and (b), the probability of an edge being chosen to trigger the cascade is proportional to its initial flow volume.

of $\delta$, each source $v$ will decrease its output by

$$\Delta = \frac{|P_v|\delta}{2n_+\langle P\rangle_+}, \tag{3.21}$$

and each sink will increase its input by

$$\Delta = \frac{|P_-|\delta}{2n_-\langle P\rangle_-}, \tag{3.22}$$

where $\langle\cdot\rangle_+$ and $\langle\cdot\rangle_-$ denote the mean source and sink strengths respectively. It is found that the same qualitative behaviour of the simplex persists in the presence of substantial noise; namely, the same inversion of $\bar{\rho}$ is observed as $q$ increases. For example, figures 3.14(a) and (b) show $\bar{\rho}$ for random Watts–Strogatz networks with $q = 0.2$, where $\xi_v$ is a random variable with mean 0 and deviation 0.2 in (a) and 0.3 in (b). These networks contain sources and sinks with normally distributed strengths, which have significant variation whilst still preserving node type in all realisations that we consider. By comparison with figure 3.10, which shows the same networks but with no noise, it is clear to see the same simplex features remain.

Figures 3.14(c) and (d) show the same networks as in (a) and (b), but where source and sink sizes are drawn from gamma distributions, parameterised by shape $k$ and scale $\theta$. Values are $(k, \theta) = (1, 2)$ and $(k, \theta) = (2, 1/2)$ in (c) and (d) respectively. Whilst the same relative behaviour within the simplex persists with gamma distributed noise, giving highest values of $\bar{\rho}$ around

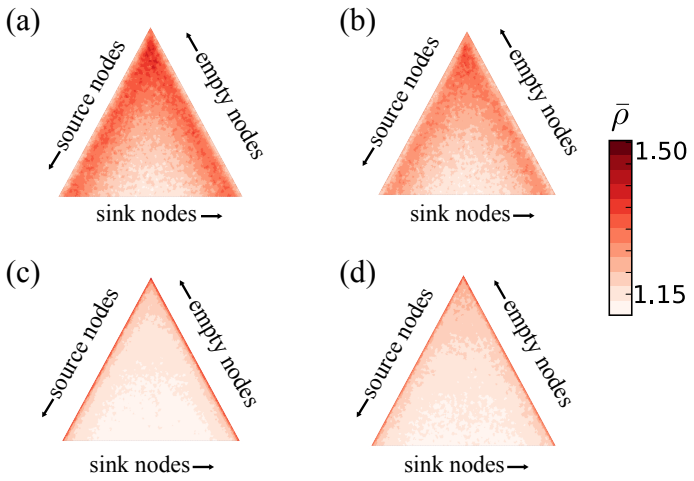



Figure 3.14: $\bar{\rho}$ for an ensembles of 200 Watts–Strogatz networks all with $q = 0.1$, $n = 100$ and $\bar{K} = 4$. In (a) and (b) the source and sink sizes are perturbed by a normally distributed random variable with mean 0 and a deviation of 0.2 and 0.3 respectively. In (c) and (d), source and sink sizes are drawn from gamma distributions with $(k, \theta) = (1, 2)$ and $(k, \theta) = (2, 1/2)$ respectively.

the perimeter for highly ordered and small world networks, the amplitudes of $\bar{\rho}$ are smaller. This indicates that skewed source-sink size distributions lend greater resilience to networks, while still preserving the same trend with respect to $q$ observed in figure 3.10.

### 3.5.2 Resilience real power grid topologies

Having established general results on synthetic networks in Section 3.5.1, we shall now proceed to look at several real world test cases of power grids. The resilience of the Austrian national grid is considered in figure 3.15. The grid topological data has been obtained from APG (2017). This network comprises $n = 67$ nodes, $m = 85$ edges and has mean degree $\bar{K} = 2.53$. Each node represents a sub-network of node-types and edges at both meso-and micro-scales that are unresolved by the data of (APG, 2017). Figure 3.15(b) shows $\bar{\rho}$ with each pixel entry averaged from 100 realisations of node-type location and with the sources and sinks assumed to have reciprocal strengths. The resulting resilience landscape is characteristic of a state intermediate of those

shown in figures 3.10 (c) and (d), where the lowest values of $\bar{\rho}$ are found in the centre of the simplex. This indicates the resilience of the Austrian power grid is enhanced by distributing power generation across many small sources. Indeed when half of the nodes are generators and the other half consumers, the resilience has a value $\bar{\rho} < 1$. Having equal numbers of sources and sinks therefore leads to a state where the grid's normal operating capacity is more than enough to cope with a cascade.

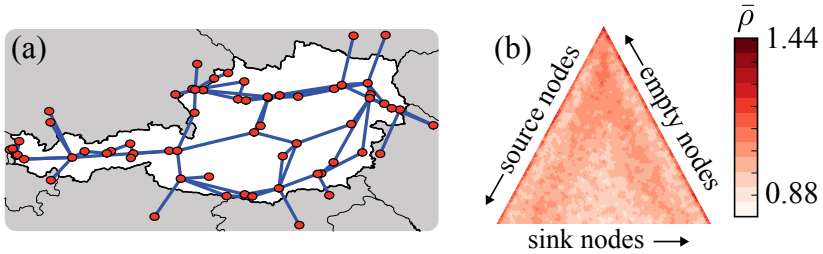



Figure 3.15: (a) Power grid topology of Austria, constructed using data from (APG, 2017). (b) $\bar{\rho}$ for this grid projected onto the node configuration simplex. Each pixel is an ensemble average over 100 random source-sink locations.

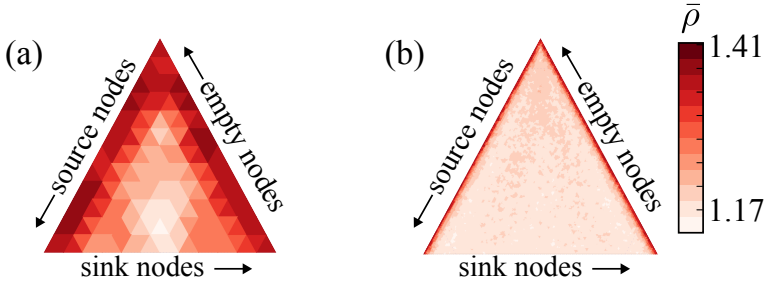



Figure 3.16: $\bar{\rho}$ in the node configuration simplex for (a) the IEEE 14 bus test networks and (b) the IEEE 188 bus test network. All pixels represent an ensemble of 50 random placements of the $(n_+, n_-, n_p)$ nodes.

As further examples on power grid topologies, figures 3.16(a) and (b) show the resilience simplex for the the IEEE 14 and IEEE 118 bus test networks. In both cases, $\bar{\rho}$ is highest around the perimeter of the simplex, resembling the case for regular and small world networks.

## 3.6 Conclusion

This chapter has considered how the resilience of networks to cascading failures is affected by the composition of node-types and topology. For the case of grids comprising a few large generators figure 3.9(a) shows that the resilience can be approximated by $\bar{K}/(\bar{K}-1)$, where $\bar{K}$ is the mean network degree. For the case of renewable energy electrical-grids, the resilience is found using simplexes of the type shown in figure 3.10. Here the appropriate node-types correspond to locations of passivity, generation or consumption of power, and these functions will mutate diurnally in response to changes in demand (Parhizi et al., 2015).

Crucially, although networks with homogeneous node-type of a given topology can be resilient to failure, this is no longer the case for a network of the same topology but with a heterogeneous composition of node-types. Regular lattices and ordered networks are most resilient when the number of node-types are similar. This changes as network's structure becomes increasingly disordered, with the least resilient configurations becoming those with similar numbers of node-types. This means that whether having only a few large sources, or many small distributed sources is most resilient depends upon how disordered the network is. Nevertheless, there still exist regimes of resilience with a range of node-type compositions that can be accessed across the entire spectrum of networks. This Chapter has relied on steady state approximations to compute flow; the subsequent Chapters investigate a fully dynamical version and reveal the impact of transient behaviour on network resilience.

# Chapter 4

# The swing equation

## 4.1 Introduction

The previous chapters have modelled network flows as steady states. Chapter 2 identified the equivalence between Nash equilibria and steady state solutions of Kirchoff's laws. This allowed the use of the Price of Anarchy as a measure of flow inefficiency. Chapter 3 assessed the resilience of networks to cascading failures, where the flow was computed directly from Kirchoff's laws. These steady state models, however, may fail to capture important dynamical properties seen in real world power grids. Transient dynamics, which can cause network failures, and the emergent power-balancing and stabilising properties of power grids (Vasquez et al., 2009), cannot be modelled using a purely steady-state approach.

The swing equation is a differential equation based model often used in power systems analysis, which is used to analyse the dynamical properties of power grids. This chapter introduces the swing equation and compares it to the steady state power flow model used previously. Following the theme of the previous two chapters, its dynamical properties will then be investigated as a function of the proportions of power flow source and sink nodes on the network. This analysis has not been conducted before, and reveals the impact of distributed generation. The ability of a power grid to function properly

will be gauged using the critical coupling capacity, which determines how readily a network will synchronise. Following chapters will then use the swing equation to reveal the resilience of fully dynamic electrical network models. This will also determine the limitations of a purely steady state approach.

Section 4.2 introduces the swing equation model and compares its steady states to those obtained from the linearised DC power model flow model used in chapter 3. Section 4.3 computes the critical coupling capacity as a function of node composition, to reveal the impact of distributed generation on the swing equation dynamics, and section 4.4 investigates the transient effects of edge removal.

## 4.2 Definition and comparison with steady state flow

This section introduces the dynamical electrical network model, and compares it to the steady state model used in chapter 3. As usual, the networks under consideration are of the form $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with $n = |\mathcal{V}|$ nodes and $m = |\mathcal{E}|$ edges. $n_+$ nodes will be chosen as sources of electrical power, $n_-$ as sinks, with the remaining $n_p$ acting as passive conduits so that $n_+ + n_- + n_p = n$. The electrical flow dynamics over the networks are modelled using the swing equation (Filatrella et al., 2008), a detailed derivation of which was presented earlier in this thesis in chapter 1, section 1.5.2. Let us briefly summarise it once again.

The swing equation relies on the fact that power grids contain rotational inertia, and models the nodes of the network as rotating machines. Each node $i$ has a mechanical phase angle

$$\phi_i(t) = \Omega t + \theta_i(t). \tag{4.1}$$

$\Omega$ is the grid reference frequency and has a value of $2\pi \times 50$Hz in Europe, most of Asia and Africa, and $2\pi \times 60$Hz in North America; $\theta_i$ is then the phase difference of node $i$. As detailed in section 1.5.2, a consideration of

power conservation at each node leads to the following equation of motion

$$\frac{\mathrm{d}^2\theta_i}{\mathrm{d}t^2} + \gamma_i \frac{\mathrm{d}\theta_i}{\mathrm{d}t} = P_i(t) + \sum_{j=1}^{n} K_{ij} \sin(\theta_i - \theta_j). \tag{4.2}$$

This is the standard form of the swing equation, as used in the electrical power systems engineering. Equation (4.2) balances dissipative power on the left, with the power generated (or consumed) $P_i$ at each node and the power transmitted from adjacent nodes, given by the summation term. The parameters are

$$\gamma_i = \frac{2D_i}{J_i}, \tag{4.3}$$

where $D_i$ and $J_i$ are the damping and moment of inertia respectively of node $i$. The parameter $\gamma_i$ is therefore the ratio of damping to inertia for each node, and has units of $\mathrm{s}^{-1}$. The power term in (4.2) is given by

$$P_i(t) = \frac{P_i^{\mathrm{in}}(t) - D_i\Omega^2}{J_i\Omega}, \tag{4.4}$$

where $P_i^{\mathrm{in}}$ is the power consumed or generated at node $i$, measured in units of Watts. $P_i$ has units of $\mathrm{s}^{-2}$. The coupling coefficient $K_{ij}$ is defined as

$$K_{ij} = \frac{A_{ij}V_0^2 B_{ij}}{J_i\Omega}, \tag{4.5}$$

where $V_0$ is the grid voltage, which is assumed to be constant throughout the network, and $B_{ij}$ is the susceptance of the edge connecting nodes $i$ and $j$. $A_{ij}$ is the $ij^{\mathrm{th}}$ element of the adjacency matrix. $K_{ij}$ also has units of $\mathrm{s}^{-2}$. For simplicity, and following Witthaut and Timme (2012), it will be assumed that damping, inertia, and susceptance are homogeneous throughout the network. This gives $\gamma_i = \gamma \quad \forall i$ and $B_{ij} = B \quad \forall ij$. The coupling coefficient can then be rewritten as $K_{ij} = \kappa A_{ij}$, where $\kappa$ is a constant. Typical values of $\gamma$, $\kappa$ and $P_i$ for a power grid are all of order $\mathcal{O}(1)$ (Manik et al., 2014). If node $i$ is a source, then $P_i > 0$; if it is a sink then $P_i < 0$. Throughout any calculations, supply and demand will be matched so that

$$\sum_{i=1}^{n} P_i = 0. \tag{4.6}$$

Equation (4.2) gives system of $n$ second order oscillators; which can be split into a system of $2n$ first order equations of the form

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \omega_i \tag{4.7}$$

$$\frac{\mathrm{d}\omega_i}{\mathrm{d}t} + \gamma\,\omega_i = P_i(t) + \kappa \sum_{j=1}^{n} A_{ij} \sin\left(\theta_i - \theta_j\right). \tag{4.8}$$

The system in equation (4.8) can then be integrated using standard time-stepping techniques. The power flow on each edge $ij \in \mathcal{E}$ is a function of the phase difference between the adjacent nodes;

$$f_{ij} = \kappa \sin\left(\theta_i - \theta_j\right). \tag{4.9}$$

In order for a power grid to function, the whole network must be synchronised to the grid's reference frequency $\Omega$. Thus a valid operating condition is satisfied when the frequency $\omega$, and its rate of change, are equal to zero (Witthaut and Timme, 2012). Such a state is a solution to the steady state equation

$$P_i - \kappa \sum_{j=1}^{n} A_{ij} \sin\left(\theta_i - \theta_j\right) = 0. \tag{4.10}$$

Finding the steady operating states in power networks with the swing equation therefore amounts to finding an $n$-dimensional vector of phase angles $\theta$ such that equation (4.10) is satisfied. In practice, during the operation of a power grid, the frequencies are allowed to drift slightly from $\Omega$. Typically, a tolerance of 1Hz drift is permitted before automatic shut-downs occur to protect the grid. The impact of this drifting and de-synchronisation will be explored in chapter 5.

It will now be demonstrated that, under reasonable assumptions, a steady state of the swing equation is a solution to Kirchoff's laws and equivalent to the steady states found using the linearised DC power flow model in chapter 3. To do this, we first rewrite equation (4.10) in vector form;

$$P - \kappa E \sin\left(E^T\theta\right) = 0, \tag{4.11}$$

where $P \in \mathbb{R}^n$ is the power vector whose $i^{\text{th}}$ entry is $P_i$, and $E$ is the $n \times m$

node-edge incidence matrix. Taking a small-angle approximation in equation (4.11) gives

$$P - \kappa\, EE^T\theta = 0, \tag{4.12}$$

and so

$$\kappa\mathcal{L}\theta = P, \tag{4.13}$$

where $\mathcal{L} = EE^T$ is the graph Laplacian matrix, also referred to as Kirchoff's matrix. The small-angle approximation requires that the phase angle difference between adjacent nodes is small; a reasonable assumption in most power grids (Filatrella et al., 2008). Equation (4.13) is now the linearised power flow equation. Solving equation (4.13), either by taking the pseudo-inverse of $\mathcal{L}$ or grounding one of the nodes to make the system non-singular, gives a direct solution to Kirchoff's laws. If the coefficient $\kappa$ is not a scalar, but a vector containing the value $\kappa_e$ for each edge $e$, then it can be incorporated into $\mathcal{L}$ to form a weighted graph Laplacian $\mathcal{L}_\kappa = E\,\mathrm{diag}(\kappa)\,E^T$.

To illustrate the equivalence between steady states of the swing equation and those obtained via direct solutions of Kirchoff's laws, figure 4.1 shows both states plotted together for two different test networks, as well as the rate of convergence. Figure 4.1(a) shows the two steady states for a ring lattice network with $n = 32$ and common degree $k = 4$. Half of the network's nodes are power sources and the other half are sinks, so that $(n_+, n_-, n_p) = (16, 16, 0)$. The total power injected into the network is $P_0 = 1$. The blue line indicates the power flow level of each edge computed via a direct solution of Kirchoff's laws, while the dashed red line is the steady state of the swing equation. The two flow patterns clearly coincide. Figure 4.1(b) shows the same experiment but for a Poisson network, also with configuration $(n_+, n_-, n_p) = (16, 16, 0)$. The equivalency of the two patterns also clearly holds for this random network. This is evidence that in the steady state, the flow patterns obtained from the swing equation are identical to those from a direct solution to Kirchoff's laws. Figures 4.1(c) and (d) show how the swing equation converges on the steady state computed via Kirchoff's laws, for the lattice and Poisson networks respectively. In both plots, the time-series of the difference between the swing equation's flow pattern $f_{\mathrm{sw}}$, and the flow pattern computed directly from Kirchoff's laws

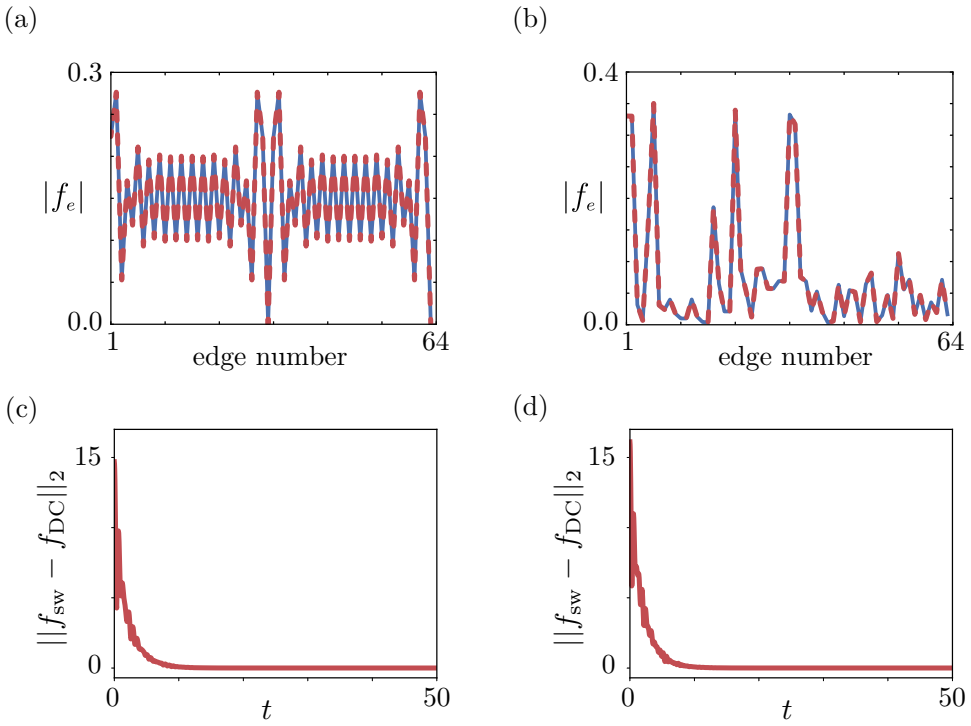



Figure 4.1: The edge flow volumes $|f_e|$ for each edge $e$ in two networks with 64 edges and 32 nodes. The ordering of the edges on the horizontal axis of each plot is arbitrary. The solid blue line in each plot is the flow pattern obtained via a direct solution to Kirchoff's laws, as given in eq. (3.1). The dashed red lines indicate the steady state flow pattern of the swing equation. Panel (a) is for a regular ring lattice and (b) is a Poisson network generated using the Watts–Strogatz method with topological randomness $q = 1$. Both networks have equal numbers of source and sink nodes, and no passive nodes. (c) and (d) show the time-series of the discrepancy between the swing flow pattern $f_{\mathrm{sw}}$ and the direct-Kirchoff $f_{\mathrm{DC}}$ states for the networks in (a) and (b), respectively. The discrepancy is measured in the 2-norm, and time $t$ is in units of seconds.

$f_{\mathrm{DC}}$, is measured in the 2-norm. In both (c) and (d), the swing equation is initialised at time $t = 0$ with a random initial state. Panel (c) shows the time-series for the lattice from figure 4.1(a), while (d) shows the case for the Poisson network from (b). In both cases, the swing equation can be seen to converge quickly to the Kirchoff steady state after approximately 5s.

## 4.3 The critical coupling strength

Although the steady state flow patterns of the swing equation are equivalent to direct solutions of Kirchoff's laws, as shown in the previous section, it is the swing equation's ability to capture non-equilibrium dynamics that has established its use in power systems engineering. These dynamic properties, such as transients and synchronisation dynamics, mean that analysis based purely on steady state methods give only an incomplete picture.

This section investigates the ability of power networks to synchronise; in particular, how that ability is impacted by the proportions of electrical power sources and sinks in the network. The aim is to determine whether having a more distributed spread of generators throughout a network allows for easier synchronistion than having only a few large generators, as in traditional power grids. This issue has not been previously addressed in the literature, and is of increasing importance as power grids become progressively more distributed. To investigate the impact of this trend in power grid design, the critical coupling value $\kappa_c$ will be computed across a range of networks with variable numbers of consumers and generators. $\kappa_c$ is the minimum edge coupling value $\kappa$ for which the network can achieve a synchronous state, as required for power grid operation. The lower the value of $\kappa_c$, the easier it is for the network to synchronise. A low value of $\kappa_c$ is therefore desirable in the design of an electrical grid as it means less excess capacity will have to be installed to ensure grid function.

Recall that the coupling coefficient is given by

$$\kappa = \frac{V_0^2 B}{J\Omega}, \tag{4.14}$$

where $B$ is the edge susceptance. Increasing $\kappa$ for an edge therefore amounts to increasing the susceptance, or, equivalently, decreasing the impedance. Note that the numerator $V_0^2 B$ in equation (4.14) is a quantity of power and has units of Watts. This value is sometimes referred to as the power rating or power capacity of an edge (Manik et al., 2014). The total power injected

into the network is given by

$$P^{\text{tot}} = \sum_{k=1}^{n_+} P_i, \tag{4.15}$$

where $k$ indexes only over source nodes, and $P_i$ is defined in equation (4.4). The quantity $\kappa/P^{\text{tot}}$ is therefore a dimensionless measure of the edge coupling strength relative to the total power in the network. Throughout this chapter, the critical coupling capacity $\kappa_c$ will be normalised in this way. Specifically, $\kappa_c/P^{\text{tot}}$ is the critical edge coupling per the total power in the network. However, in this section a value of $P^{\text{tot}} = 1$ is used so that $\kappa_c/P^{\text{tot}} = \kappa_c$. For notational brevity, from here on $\kappa_c$ is to be understood as the dimensionless quantity, normalised by $P^{\text{tot}} = 1$.

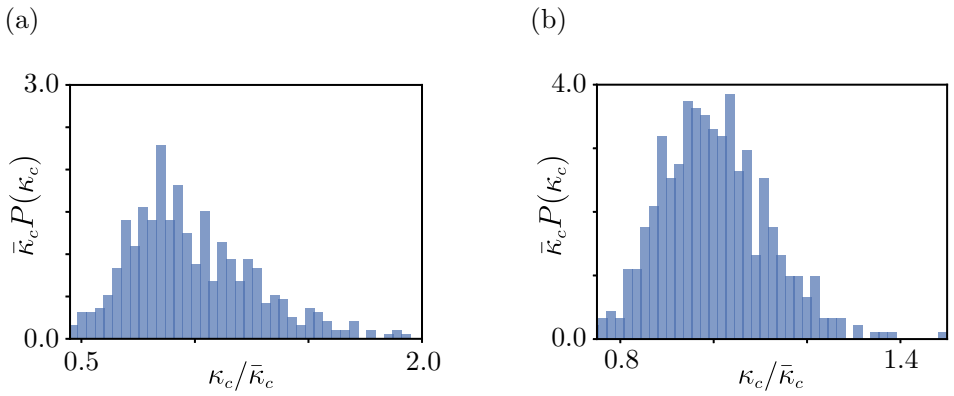



Figure 4.2: Distributions of critical coupling strengths $\kappa_c$ for: (a) an ensemble of 200 small world networks with $n = 50$, $q = 0.1$ and $(n_+, n_-, n_p) = (30, 30, 0)$, and (b) a Poisson network also with $n = 50$ and $(n_+, n_-, n_p) = (30, 30, 0)$.

For simple one and two node systems $\kappa_c$ can be found analytically (Witthaut and Timme, 2012); however, no such analysis exists for the general case of complex networks of arbitrary size. The dependence of $\kappa_c$ on network node composition will therefore be computed numerically as follows. First, for a given network with node composition $(n_+, n_-, n_p)$, the initial steady state of the swing equation will be found by numerically integrating equations (4.8) for a high value of $\kappa$. Typically a value of $\kappa \sim 5$ will be used to ensure a synchronous state is found. Then, the value of $\kappa$ will be gradually

reduced using continuation until a synchronous state can no longer be found. This point is the critical value $\kappa_c$. This process will be repeated across an ensemble of random networks, with identical node configuration $(n_+, n_-, n_p)$, to produce a distribution of $\kappa_c$. Two examples of such distributions are shown in figure 4.2. The mean critical coupling capacity $\bar{\kappa}_c$ is then computed from the ensemble. This process is then repeated for each of the $n(n-1)/2$ possible configurations of $(n_+, n_-, n_p)$, and each value projected into the node configuration simplex as used in chapters 2 and 3. This will reveal how $\bar{\kappa}_c$ varies as a function of node composition for each specific network type.

Note that there are likely to be many possible probability distribution functions that could provide a good fit for the data in figure 4.2. However, since there is presently no working hypothesis that should lead us to favour any particular distribution over another, the data shall be left un-fitted so as not to imply any spurious underlying physical explanations.

Figures 4.3(a) to (d) show $\bar{\kappa}_c$ plotted in the node configuration simplex for ensembles of networks constructed using the Watts-Strogatz method (Watts and Strogatz, 1998), where all networks have size $n = 50$, initial degree $K = 4$, random rewiring parameter $q$, and $m = 100$. Each pixel in the simplexes is obtained from distributions of 200 such network realisations. Panel (a) shows the case for regular lattices with $q = 0$. The highest value of $\bar{\kappa}_c$ is located at the apex of the simplex, representing the case where the networks have only one source and one sink, with all others passive. This means that such a node configuration offers the worst possible synchronising ability. Down the lateral edges of the simplex, representing networks with only a few big sources or sinks, the values of $\bar{\kappa}_c$ remain high. The lowest values of $\bar{\kappa}_c$, and thus the networks which can synchronise most readily, are found in the interior of the simplex. This interior region of the simplex represents networks with a high level of distribution, possessing many small source and sink nodes. The absolute lowest value of $\bar{\kappa}_c$ is found at the bottom centre of the simplex, representing networks with an equal number of sources and sinks, and no passive nodes. These observations support the hypothesis that increasingly distributed generation leads to networks that can synchronise more easily with lower values of coupling. The same trend is also seen for the

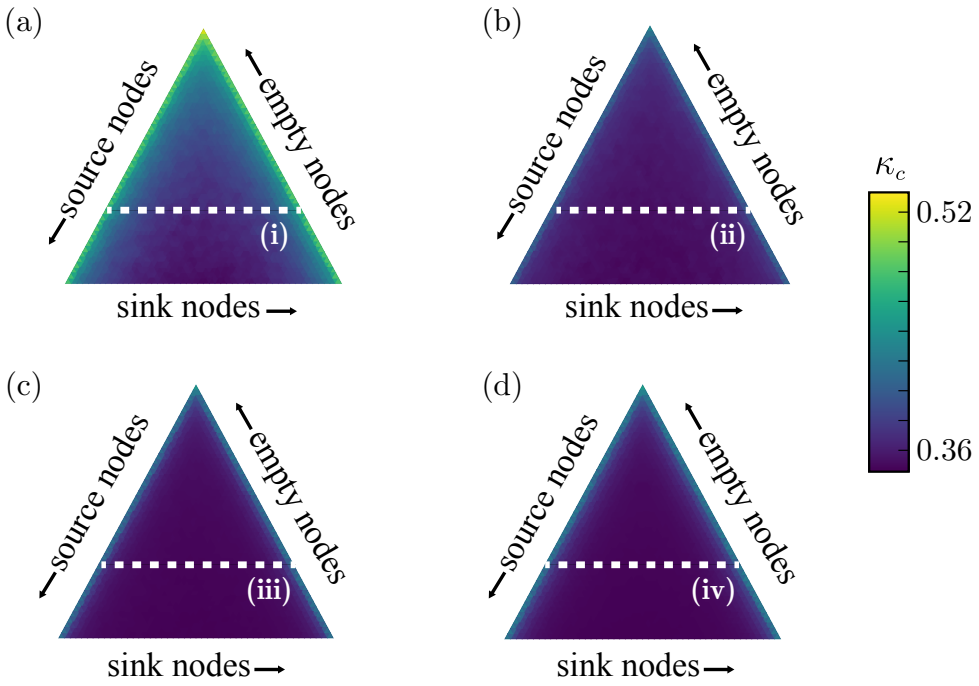



Figure 4.3: Simplex plots showing the variation of the mean coupling capacity $\bar{\kappa}_c$ as a function of the numbers of source $n_+$, sink $n_-$, and passive or empty $n_p$ nodes. Each point on each simplex represents a unique configuration of $(n_+, n_-, n_p)$, and the value of each point is determined from an ensemble average over 200 networks. All networks have $n = 50$ nodes and $m = 100$ edges. Simplex (a) shows the variation of $\bar{\kappa}_c$ for regular ring lattices. (b), (c) and (d) show the cases for Watts-Strogatz networks with random rewiring parameter $q$ values of 0.1, 0.4 and 1.0 respectively. Each of the slices (i) to (iv) are plotted separately in figure 4.4.

small-world networks with $q = 0.1$ in figure 4.3(b), and the Poisson networks with $q = 0.4$ and $q = 1.0$ in 4.3(c) and 4.3(d) respectively. The amplitudes of $\bar{\kappa}_c$ in these more random networks are also uniformly lower than for the lattice networks. This indicates that for all node configurations, random networks synchronise more readily than lattices. These features can be seen clearly in figure 4.4, which shows slices through each of the four simplexes in figure 4.3. These cross-sections demonstrate that $\bar{\kappa}_c$ is minimised when there are equal numbers of sources and sinks.

For a real world demonstration, $\bar{\kappa}_c$ shall now be computed as a function of the node configuration for the Austrian national grid, as used previously in chapter 4. This network, illustrated in figure 4.5(a), is constructed from

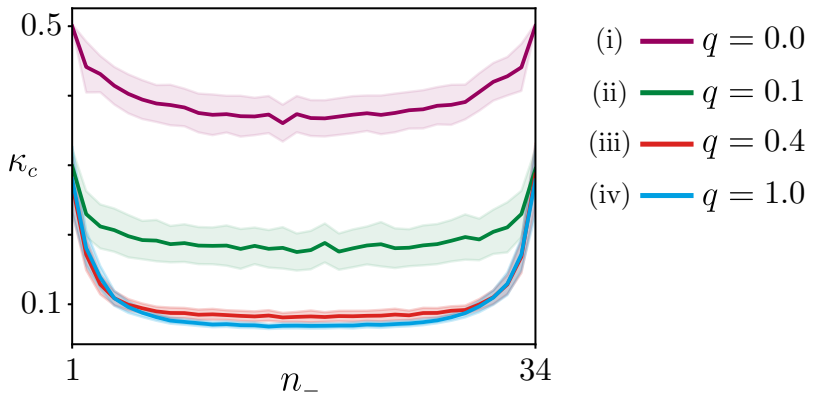



Figure 4.4: Each of the slices (i)-(iv) from figure 4.3 plotted individually. Each line represents the critical coupling $\bar{\kappa}_c$ as a function of the number of sink nodes, for a fixed number of source nodes $n_+ = 15$. Each data point is obtained from an average of 200 network realisations; the shaded regions indicate standard deviation. In each case $\bar{\kappa}_c$ is minimised in the centre, indicating the point at which there is an equal number of sources and sinks. As randomness $q$ increases, the values of $\bar{\kappa}_c$ decrease uniformly.

data from (APG, 2017) and contains 67 nodes and 85 edges. Each node is a substation, behind which is either a load (net consumer) or a generator. Each of the edges is a transmission line. As before, the damping and inertia of the nodes will be assumed to be constant so that $\alpha_i = \alpha$ for all nodes. 4.5(b) shows the normalised critical coupling $\bar{\kappa}_c$ for the Austrian grid and demonstrates that the lowest values are found in the centre of the simplex, just as in the case for the synthetic random networks of figure 4.3. A slice through the centre of the simplex, plotted separately in panel 4.5(c), confirms this and shows that $\bar{\kappa}_c$ is minimised when the numbers of sources and sinks is equal. This shows that more distributed generation allows the Austrian grid to synchronise more easily, making it more robust against perturbations.

## 4.4 Impact of network edge removals

A small perturbation to a stable steady state of the swing equation will decay back to a synchronous state. Such a perturbation is referred to as a transient.

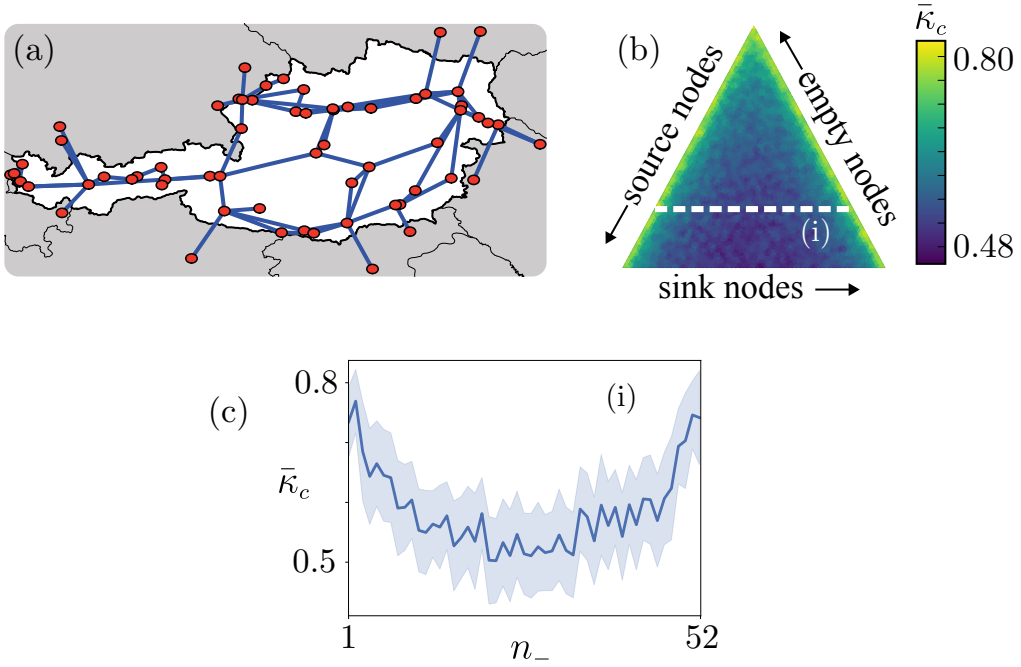



Figure 4.5: (a) The Austrian power grid constructed using data from (APG, 2017). (b) The mean critical coupling capacity $\bar{\kappa}_c$ as a function of node configuration for the Austrian grid. (c) The cross-section (i) from panel (b) plotted separately, demonstrating that $\bar{\kappa}_c$ is minimised in the centre of the simplex where $n_+ = n_-$.

Analysing the amplitude and duration of transients is a common form of stability analysis in power systems engineering. The reader might refer to Chiang et al. (1987) for an overview.

This section investigates transients that are caused by the sudden removal of a single edge from the network, and investigates how their severity is influenced by the inertia and damping on the grid, as well as by the flow volume on the removed edge. If the networks possess significant transient behaviour, this may be critical to the operation of the network. For instance, if the frequency drifts too far from $\Omega$ during the transient, the network may begin to shut down as frequency monitoring devices are triggered. Additionally, edge flow capacities may be tripped during transients, significantly changing the course of a cascading failure.

All transients in this section are simulated by first finding a synchronous

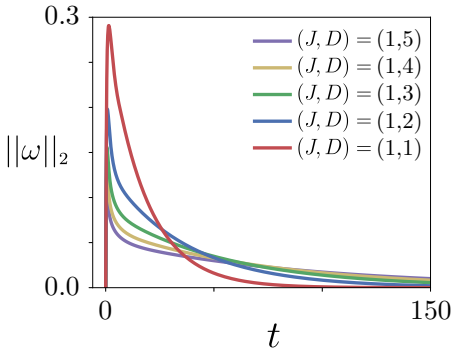



Figure 4.6: The amplitude of transients as a function of time $t$ in seconds in a regular lattice with common degree $K = 4$ and configuration $(n_+, n_-, n_p) = (25, 25, 0)$. Transients are shown for various different values of the parameter $\alpha = 2D/J$.

state of the network by integrating equation (4.8), and then deleting the most heavily loaded edge. The amplitude and duration of the resulting transient are then recorded. Figure 4.6 shows such transients for a regular lattice with $n = 50$ nodes, common degree $K = 4$, and node configuration $(n_+, n_-, n_p) = (25, 25, 0)$. The amplitudes of the transients in figure 4.6 are measured by the 2-norm $||\cdot||_2$ of the $n$-dimensional vector $\omega$, whose $i^{\text{th}}$ element is $\omega_i$. A value of $||\omega||_2 = 0$ indicates that the network is synchronised to the grid frequency $\Omega$, as is the case before the edge is removed at time $t = 0$. Figure 4.6 reveals that for all values of $\gamma$ that are tested, the amplitude of the transients never exceeds 0.3. This indicates that their amplitudes are small and within the tolerance of a typical grid. Frequencies may typically be allowed to drift up to around 1Hz away from $\Omega$ before shutdowns are triggered. As $\gamma$ is increased, meaning that the damping increases relative to the inertia on the grid, the amplitudes decrease yet further. Nonetheless, the durations of these small transients are very long. For $\gamma = 1/2$, the network returns to synchrony after about $t = 60s$. Higher values of $\gamma$ result in small transients lasting more than 2 minutes.

The transients in figure 4.6 were triggered by removing the most heavily loaded edge. We now turn our attention to how the choice of triggering

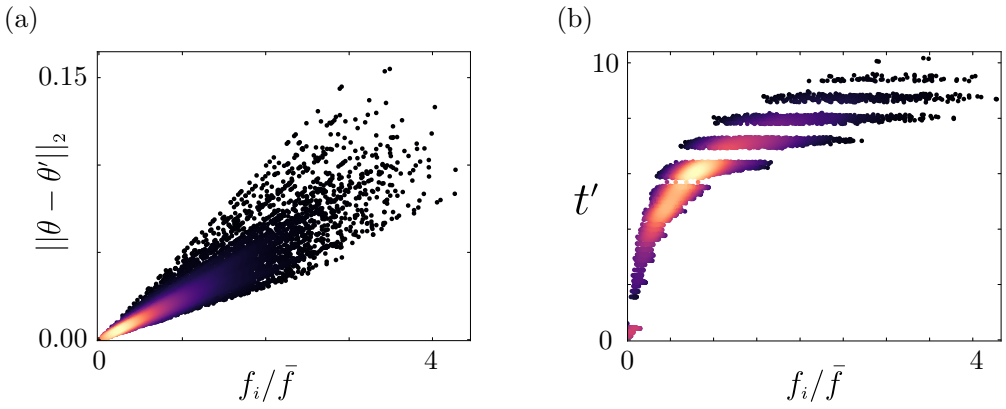



Figure 4.7: (a) The amplitude of the transients, measured by the deviation from the steady state $||\theta - \theta'||_2$, as a function of the relative flow volume on the triggering edge for ensembles of lattices with $K = 4$ and $(n_+, n_-, n_p) = (10, 40, 0)$. (b) The length of time taken for the transients to drop beneath $||\omega||_2$ on the same networks as in (a). Brighter colours here denote higher density of data points.

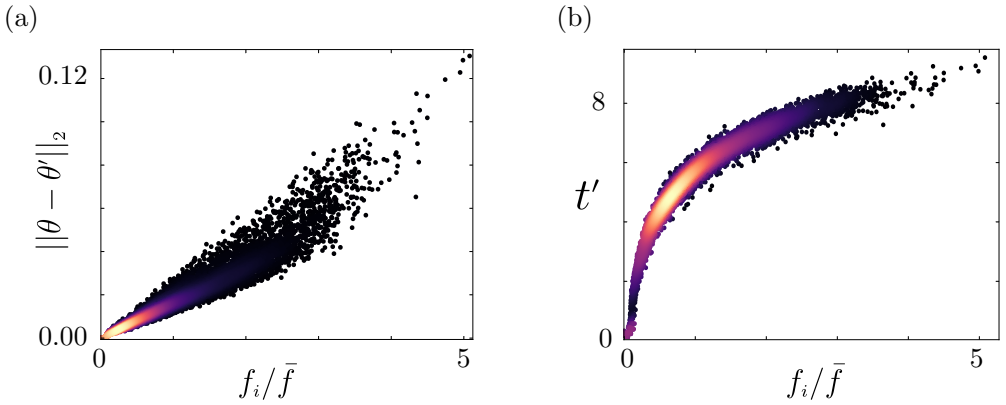



Figure 4.8: (a) The amplitude of the transients, as a function of the relative flow volume on the triggering edge for ensembles of Poisson networks with $q = 1$, $K = 4$ and $(n_+, n_-, n_p) = (10, 40, 0)$. (b) The length of time taken for the transients to drop beneath $||\omega||_2$ on the same networks as in (a).

edge effects the transients. This is investigated using an ensemble process as follows. First, an ensemble of random Watts-Strogatz networks is instantiated, each with identical parameters $q$, $K$, $n$ and $(n_+, n_-, n_p)$ but with source-sink node locations chosen at random. For each network, a steady state

is found and then a transient triggered by removing a random edge. The flow volume on that edge, relative to mean flow volume in the network is recorded. This is denoted $f_i/\bar{f}$. The maximum deviation caused by the resulting transient, measured as $||\theta - \theta'||_2$ is recorded. Here $\theta$ denotes the steady state solution, and $\theta'$ is the state at the maximum of the transient. The time taken for the frequency $||\omega||_2$ to drop back beneath 0.1 is also recorded, and denoted $t'$. The quantities $||\theta - \theta'||_2$ and $\theta'$ are then plotted against relative load $f_i/\bar{f}$ of the triggering edge to reveal that impacts the severity of the transients. Figure 4.7 shows such an experiment of a regular lattice with $q = 0$, $K = 4$ and $(n_+, n_-, n_p) = (10, 40, 0)$. The size of the transient increases linearly with the flow volume of the triggering edge. The duration also increases, but plateaus off for very heavily loaded edges. The same trends are also observed in the case of Poisson networks with $q = 1.0$, $K = 4$ and $(n_+, n_-, n_p) = (10, 40, 0)$, as shown in figure 4.8. These results show clearly that the severity of the transients correlate with the flow load on the deleted edge.

The linear relationship between the size of the transient and the flow volume of the triggering edge, as seen in figures 4.7(a) and 4.8(a), can be explained as follows. First, note that the flow volume $f_i$ on the triggering edge must be reapportioned across the remaining edges after it has been deleted. This linear perturbation in edge flow volumes will cause a linear change in the nodal values $\theta$. This is because the relationship between edge flows and node values, given in equation (4.9), is approximately linear under the realistic small-angle assumption.

## 4.5 Conclusion

This chapter has introduced the swing equation, a dynamical model for computing the power flow in electrical networks. The swing equation is capable of capturing dynamic properties of an electrical grid which a steady state model, such as those used in previous chapters, cannot. Although section 4.2 showed that the steady states of the swing equation are equivalent to those found from direct solutions to Kirchoff's laws, section 4.4 found evidence of

significant transient behaviour occurring between steady states. Although these transients are small in amplitude, their duration is long, particularly compared to the time-scales over which cascading failures on an electrical grid occur. This raises the possibility that much of a cascade may happen during transient behaviour in the grid, meaning that the swing equation may produce significantly different results if used to compute cascading failures, compared to the steady state model used in chapter 3. This will be investigated in detail in the next chapter.

This chapter has also investigated how the proportions of sources and sinks, or generators and consumers, on the network impact its ability to synchronise. It has been shown that for all networks, including a real-world power grid topology, synchronisation can be achieved more readily when networks contain many small sources and sinks distributed throughout their structure. This is an encouraging result, indicating that the modern trend towards increased distribution of power generation is beneficial for the dynamical properties of the network.

# Chapter 5

# Resilience of dynamic flow networks

## 5.1 Introduction

Power grids are susceptible to a variety of failures, such as overloading lines, de-synchronisation, and voltage instability (D'Souza, 2017). Chapter 3 used a steady state model to assess the resilience of power grids to cascading failures of overloading lines, showing that increasing the proliferation of small generators boosts resilience for more ordered networks and small-world networks. However, chapter 4 used a dynamical model of electrical power flow called the swing equation to show that significant transient behaviour occurs in power networks, which is not captured by a steady state model.

This chapter investigates cascading failures using the swing equation to reveal the importance of dynamics in determining resilience. Following the theme of preceding chapters, the dependence of resilience upon the numbers and proportions of generators and consumers is then investigated. This dependence is important to understand given the increasing trend towards small, distributed renewable sources of power in modern grids. Uncovering the impact of distribution will offer insight into how the proportions and volume of generation can be controlled to boost operational resilience.

Section 5.2 introduces the dynamic cascading failure model which will be used to asess the resilience of power grids. Section 5.3 then investigates how this resilience is distributed over ensembles of networks, and determines the effect of increased the mean degree. Section 5.4 then investigates the resilience as a function of node configuration, ascertaining the impact of distributed generation.

## 5.2 Dynamic cascade model

This section gives an overview of the model which will be used to simulate cascading failures on power grids. A power network's ability to survive such a failure will be used as a measure of its resilience. The model is similar to that used in chapter 3; however, the power flow will be computed using the swing equation rather than the steady state model. This will result in cascading failures that take into account transient behaviour and synchronisation dynamics, as well as overloading line failures.

As in chapters 2 and 3, all networks will be of the form $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with $n = |\mathcal{V}|$ nodes and $m = |\mathcal{E}|$ edges. $n_+$ nodes are chosen to be generators, $n_-$ to be consumers, and the remaining $n_p$ to be passive or empty such that $n_+ + n_- + n_p = n$. Generators and consumers may sometimes be referred to as source and sink nodes. The power flow over each network will be computed via the swing equation, which was examined in detail in chapter 1 and again in chapter 4. However, let us briefly restate it here for convenience. Each node $i \in \mathcal{V}$ has a phase angle $\theta_i \in [0, 2\pi)$ and a rotational frequency $\omega_i = \dot{\theta}_i$, which is measured relative to the power grid's reference frequency $\Omega$. The dynamics of $\theta_i$ and $\omega_i$ for each node are then given by the two coupled equations

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \omega_i \tag{5.1}$$

$$\frac{\mathrm{d}\omega_i}{\mathrm{d}t} + \gamma\,\omega_i = P_i(t) + \kappa \sum_{j=1}^{n} A_{ij} \sin\left(\theta_i - \theta_j\right), \tag{5.2}$$

resulting in a system of $2n$ ordinary differential equations to describe the

dynamics of the network. $A_{ij}$ denotes the $ij^{\text{th}}$ component of the adjacency matrix $A$. The parameter $\gamma$ describes the ratio of the damping $D$ of each node and its moment of inertia $J$:

$$\gamma = \frac{2D}{J}. \tag{5.3}$$

The coupling parameter $\kappa$ incorporates the susceptibility $B$ of each edge and the grid voltage $V_0$:

$$\kappa = \frac{V_0^2 B}{J\Omega}. \tag{5.4}$$

Following Witthaut and Timme (2012) the susceptibility, damping, and inertia are assumed to be homogeneous constants across the network, which will allow for a tighter focus on the specific impact of consumer and generator numbers on resilience. As in Witthaut and Timme (2015), parameter values are chosen to be $\gamma = 1$ and $\kappa = 5$. These choices are within a realistic range for a power grid and, as shown in section 4.3, and ensures that a stable synchronous state is found. The power term $P_i$ in equation 5.2 will have a value of $P_i > 0$ if node $i$ is a generator, or $P_i < 0$ if $i$ is a consumer. $P_i$ is itself given by

$$P_i = \frac{P_i^{\text{in}} - D\Omega^2}{J\Omega}, \tag{5.5}$$

where $P_i^{\text{in}}$ is the power in Watts injected into node $i$. The quantities $\gamma$, $\kappa$ and $P_i$ have units of $\text{s}^{-1}$, $\text{s}^{-2}$, and $\text{s}^{-2}$ respectively with typical values of $P_i$ being $\mathcal{O}(1)$. The power flow on each edge $ij \in \mathcal{E}$ is given by

$$f_{ij} = \kappa \sin\left(\theta_i - \theta_j\right). \tag{5.6}$$

A steady state of the system is given by the case where $\omega$ and $\dot{\omega}$ are both equal to zero, meaning that the grid is synchronised to the reference frequency $\Omega$. This is true when

$$P_i - \kappa \sum_{j=1}^{n} A_{ij} \sin\left(\theta_i - \theta_j\right) = 0, \qquad \forall i \in \mathcal{V}. \tag{5.7}$$

Steady states satisfying equation (5.7) can be obtained by time-stepping system (5.2).

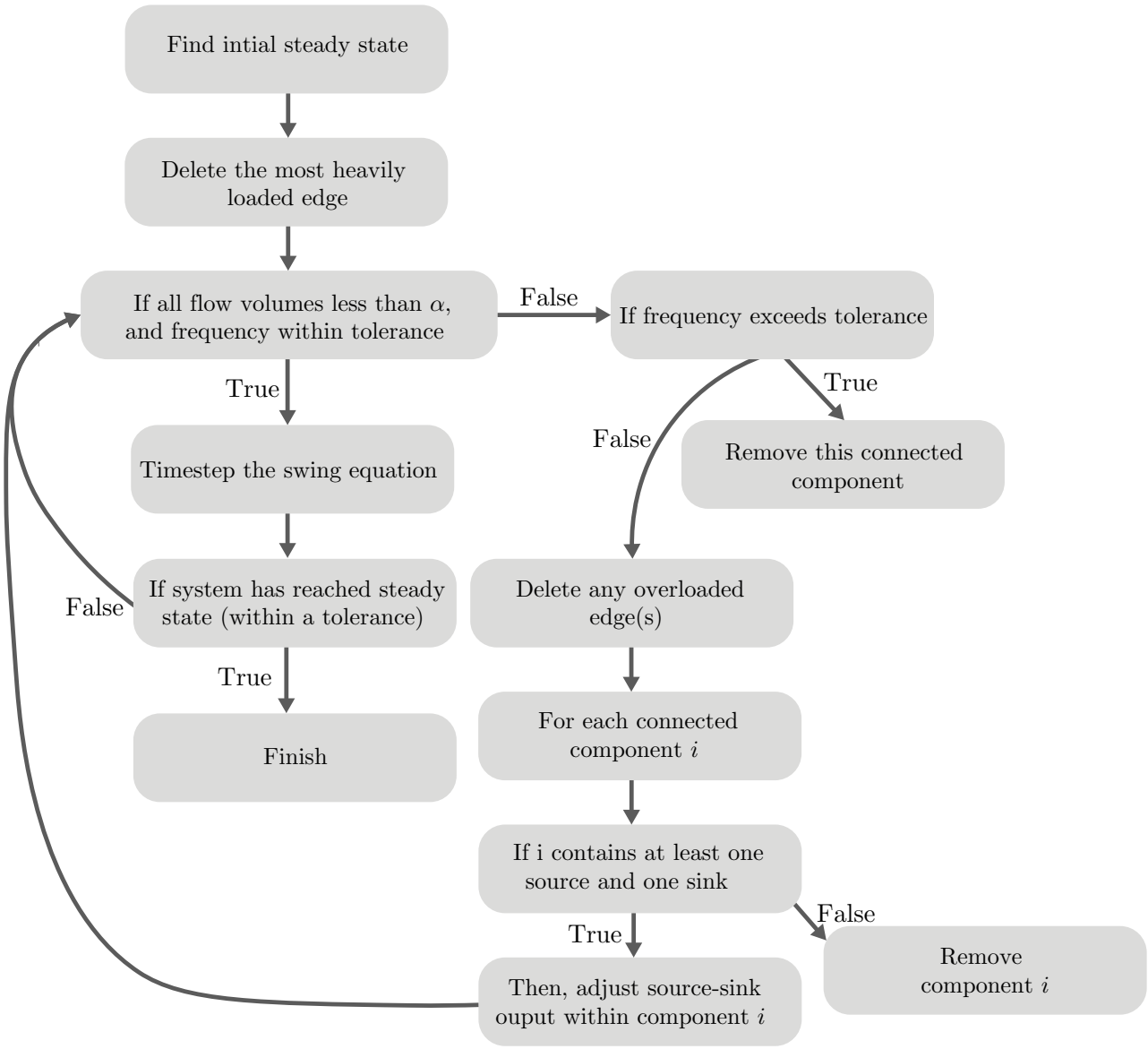



Figure 5.1: Overview of the dynamical network cascade model. The initial steady state and all power flow patterns are computed by time-stepping equation (5.2). A steady state of the system will satisfy equation (5.7); the cascade finishes when such a state has been reached (within some specified tolerance) for all network components, or when no components remain.

Having now outlined the dynamical system used to model electrical power flow, the scheme used to simulate cascading failures taking into account both line overloads and de-synchronisation proceeds as follows. Firstly, a network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is instantiated. An initial stable power flow pattern $f^{\text{init}} \in \mathbb{R}^m$ is then found by time-stepping system (5.2). A maximum flow capacity $\alpha$

is then assigned to each edge. If, at any point, the flow volume on an edge exceeds $\alpha$, then that edge will be deemed overloaded and removed from the network. The initial maximum edge flow volume is recorded and labelled $\alpha_*$, so that

$$\alpha_* := \max_{(ij)\in\mathcal{E}} |f^{\text{init}}_{(ij)}|. \tag{5.8}$$

This shall be referred to as the network's operating capacity: the minimum capacity required in order for the network to function in its initial state. The cascade is then triggered by removing the most heavily loaded edge, serving as a model for the overloading power line events observed in power grids. The power flow on the network will then naturally readjust, possibly to a new stable operating condition. The evolution of the flow is captured by continuously time-stepping the swing equation throughout the cascade process. If, after the initial edge removal, the flow volume on any other edges exceed $\alpha$, then they too are removed. The network is continuously monitored to see if it fractures into separate connected components. If such fractures occur, then the cascade process is continued in each new connected component. In each component, power generation and consumption must be matched so that

$$\sum_{i\in\mathcal{V}'} P_i = 0, \tag{5.9}$$

where $\mathcal{V}'$ is the component's node set. This matching is done by adjusting the strength of each generator and consumer node. For simplicity these adjustments are done reciprocally, as in chapter 3. Specifically, if a given connected component contains $n'_+$ generators, $n'_-$ consumers, and has a flow surplus of $\delta$, then each generator node decreases its output by $\delta/2n'_+$ and each consumer node increases its demand by $\delta/2n'_-$. If a component contains either no generators or consumers, then flow cannot be balanced and the entire component must be removed.

The frequency values $\omega_i$ are also continuously monitored to ensure the network does not drift too far from the reference frequency $\Omega$. Real-world power grids begin to shut down if the frequencies drift outside of a small tolerance, typically 1Hz, of $\Omega$ in order to prevent physical damage to hardware. To simulate this in the present cascade model, components of the networks will therefore be removed if local values of $\omega_i$ become greater that 1Hz.

This process continues, with edges being pruned if their flow volumes exceed $\alpha$, power supply being balanced where possible, and components of the grid being removed if they drift too far from $\Omega$. Eventually the system will reach a steady state where any remaining network components possess a steady electrical power flow satisfying equation (5.7) within some tolerance. In particular a tolerance of $10^{-5}$ is used for all subsequent simulations. The fraction of edges $\mathcal{S}$ surviving the cascade will then be recorded. A value of $\mathcal{S} = 0$ indicates total network failure. $\mathcal{S} = 1$ indicates complete survival and resilience to cascades. The cascade scheme is summarised in the flowchart in figure 5.1.

As in chapter 3, the value of $\mathcal{S}$ will then be determined as a function of the network's capacity $\alpha$. This is done by repeating the above cascade scheme for a range of values of $\alpha$ for each network. If the network can survive the cascade with an amount of capacity $\alpha$ less than the operating capacity $\alpha_*$, then this indicates that the network is inherently resilient to such failures. If a network only survives if it possesses a large amount of $\alpha$ in excess of $\alpha_*$, then it is highly susceptible to failure. This is because a large amount of capacity must be installed, in excess of the $\alpha_*$ required for normal operation, in order to protect it from cascading failure. Defining $\alpha_c$ as the value of $\alpha$ at which half of the network survives, i.e. $\mathcal{S} = 1/2$, suggests the resilience measure

$$\rho = \frac{\alpha_c}{\alpha_*}. \tag{5.10}$$

This is the ratio of the volume of capacity required for at least half the network to survive, to the volume of capacity required for normal operation. The smaller the value of $\rho$, the more inherently resilient the network. This is the same metric used in chapter 3, thus allowing direct comparison with the steady state model.

The cascade model used in chapter 3 relied on a steady state power flow model. The resulting cascades were therefore a sequence of steady states and were unable to capture any transient dynamics. The cascade model presented in this chapter is a dynamical system, which takes into account transient behaviour and de-synchronisation events. The rest of the chapter will use this dynamical cascade model and the resilience measure $\rho$ to investigate

the resilience of dynamic power flow networks. These results will also be compared with the results obtained from the steady state case in chapter 3.

## 5.3 Cascades on random networks

This section investigates the resilience of ensembles of random networks to the dynamical cascades described in section 5.2. Random network models will form the substrate on which general conclusions regarding resilience of complex networks can be made. Of particular interest in this section is how resilience varies across ensembles, and how long cascades take to unfold in each network.

### 5.3.1 Distribution of resiliences $\rho$

Each ensemble will comprise random networks generated using the Watts-Strogatz method (Watts and Strogatz, 1998). Every network in a given ensemble will have the same number of nodes $n$ and edges $m$, as well as identical values of topological randomness $q$ and mean degree $K$. The networks in a given ensemble will also all have the same node configuration $(n_+, n_-, n_p)$, but with the positions of the consumers and generators randomly allocated. For each network, the fraction $\mathcal{S}$ of edges surviving a cascade will be computed as a function of the network's capacity $\alpha$. From this, the resilience measurement $\rho$ is determined. $\rho$ is the ratio of the critical capacity $\alpha_c$ required to survive a cascade to the operating capacity $\alpha_*$, which is required for normal operation. The ensemble averaged fraction $\bar{\mathcal{S}}$ of surviving edges can then be computed as a function of network capacity, as can the ensemble average operating capacity $\bar{\alpha}_*$, ensemble average critical capacity $\bar{\alpha}_c$, and enesemble average resilience $\bar{\rho}$. This resilience measure $\bar{\rho}$ gives the value of capacity at which, on average, at least half a network will survive a cascade. Figure 5.2 shows these measurements for four different ensembles. Each ensemble contains 500 networks of size $n = 60$ and $m = 120$. The panel in figure 5.2(a) shows $\bar{\mathcal{S}}$ as a function of $\alpha$ for an ensemble with $(n_+, n_-, n_p) = (15, 45, 0)$

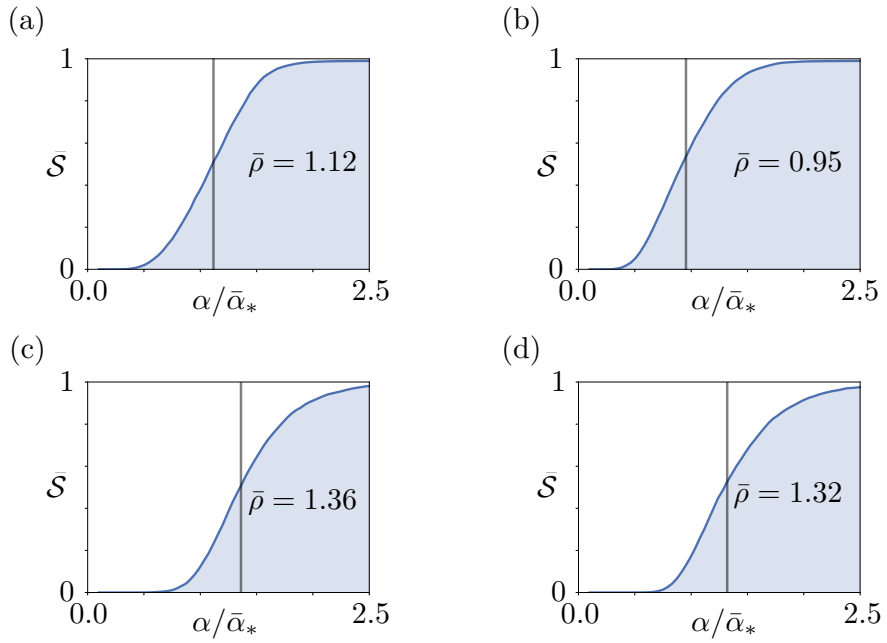



Figure 5.2: The mean fraction $\bar{\mathcal{S}}$ of edges surviving a cascade as a function of edge capacity $\alpha$ for four different ensembles of Watts-Strogatz networks. Each ensemble contains 500 network realisations. The horizontal axes have been normalised by each respective ensemble's mean operating capacity $\bar{\alpha}_*$. The grey vertical line in each plot indicates the mean resilience measure $\bar{\rho}$ for each ensemble. Panel (a) shows the case for networks with node configuration $(n_+, n_-, n_p) = (15, 45, 0)$ and $q = 0.1$. (b) shows $(n_+, n_-, n_p) = (30, 30, 0)$ and $q = 0.1$; (c) shows $(n_+, n_-, n_p) = (15, 45, 0)$ and $q = 0.6$; (d) shows $(n_+, n_-, n_p) = (30, 30, 0)$ and $q = 0.6$.

and $q = 0.1$, corresponding to networks with a small-world structure and a proportion of generators to consumers resembling a traditional power grid. In this case, $\bar{\mathcal{S}} = 0$ for small values of $\alpha$, rises sharply around the critical value, and saturates to 1 for high $\alpha$. This behaviour is repeated in the other three ensembles demonstrated in figure 5.2, albeit the position of $\bar{\rho}$ varies between them. The variation of $\bar{\rho}$ with node configuration and $q$ will be the subject of section 5.4. It can already be seen in figure 5.2 that increasing randomness $q$ leads to higher values of $\bar{\rho}$ and thus lower resilience. Additionally, panels (b) and (d) have slightly lower values of $\bar{\rho}$ than (a) and (b), indicating that increasing the level of distributed generation slightly increases resilience in these cases. The sigmoidal character of the $\bar{\mathcal{S}}$ profiles in figure 5.2 was also observed for the equivalent steady state cases in chapter 3.

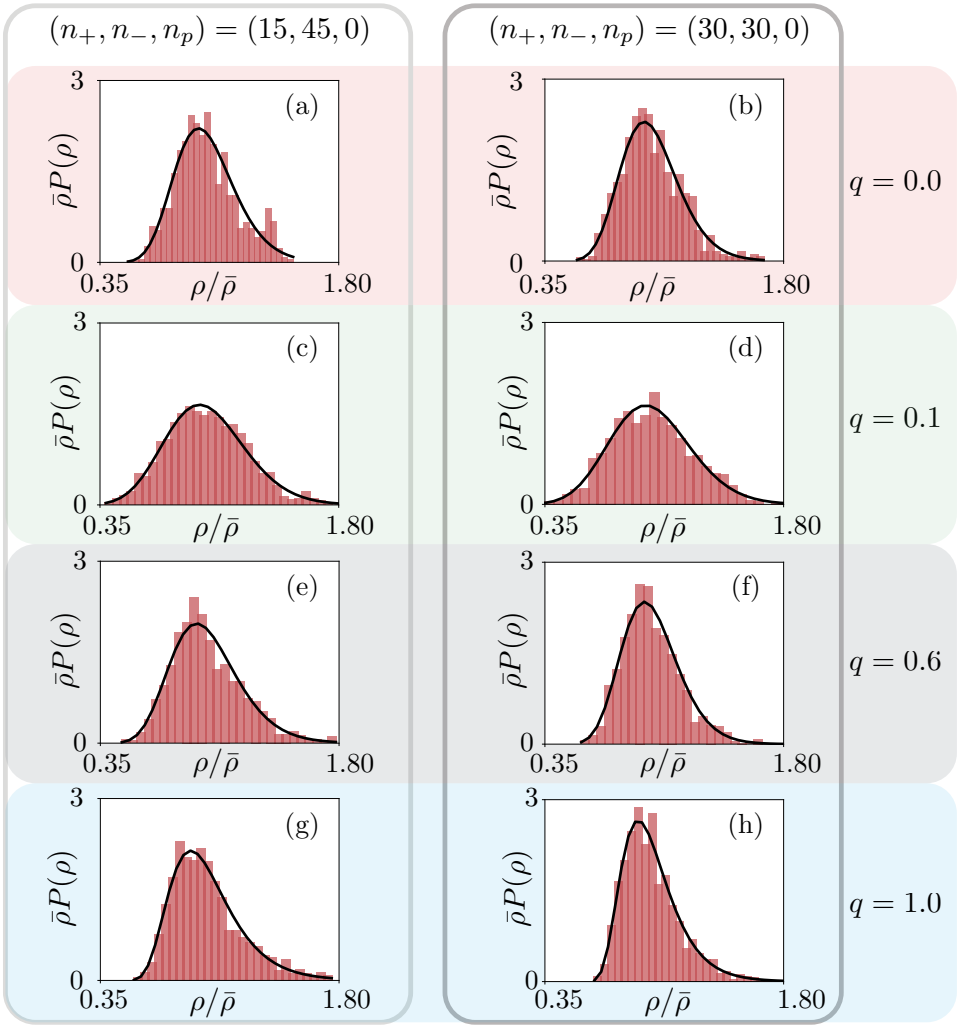



Figure 5.3: Distributions of resilience $\rho$ across 8 different ensembles, each containing 500 Watts–Strogatz networks. All horizontal axes have been scaled by the mean $\bar{\rho}$. The left-hand column of distributions is for ensembles with a node configuration characteristic of traditional grids with a few large generators. The right-hand column is for networks with equal numbers of consumers and generators, indicative of a modern distributed grid. Each of the four rows of panels represents a different value of randomness $q$, from regular lattices at the top to Poisson networks at the bottom. Each histogram is fitted to a log-normal distribution, parameterised by the mean $\mu$ and variance $\nu$. Values of $(\mu, \nu)$ are: (a) $(0.88, 0.19)$, (b) $(0.77, 0.18)$, (c) $(1.02, 0.12)$, (d) $(0.86, 0.1)$, (e) $(1.37, 0.21)$, (f) $(1.29, 0.18)$, (g) $(1.45, 0.35)$, (h) $(1.35, 0.36)$. All confirmed with Kolmogorov-Smirnov tests (see Frank and Massey, 1951) returning KS measures $< 0.03$ and $p$-values $> 0.2$, where the null-hypothesis is that the data is the gamma-distribution, indicating a good fit.

While figure 5.2 shows the variation of the ensemble mean values $\bar{\mathcal{S}}$ and $\bar{\rho}$, the individual values of $\rho$ within each ensemble are realisations of a random variable. The distribution of $\rho$ for various ensembles is shown in figure 5.3. The figure shows the case for networks with configuration $(n_+, n_-, n_p) = (15, 45, 0)$ as well as the case for networks with $(n_+, n_-, n_p) = (30, 30, 0)$, representing grids with a large numbers of small sources. The figure also shows the effect of increasing topological randomness $q$ for these ensembles. For every case tested, the distributions can be fitted to a log-normal distribution. These distributions are parameterised by their mean $\mu$ and variance $\nu$. Recall from Chapter 3 that the resilience values for steady-state cascades are also log-normally distributed. Here, as in Chapter 3, we note that this log-normality may be due to a relationship between cascades and fracturing processes. However, this relationship will not be investigated in this thesis, though it may form an interesting point for further work. Our purpose here is only to fit a convenient distribution to allow for efficient sampling of the data later on in Chapter 6.

The fit of each distribution was validated with a Kolmogorov–Smirnov (KS) test (see Frank and Massey, 1951, for an overview), which measures the maximum distance between the cumulative distribution of the data and the cumulative test distribution. These tests are performed using SciPy's KS test routine (Virtanen et al., 2020). This routine returns a KS measure: a value which, if less than approximately 0.5, means that the data is likely to come from the test distribution. In the case of figure 5.3, all KS measures are less than 0.03. Since these KS measures are somewhat hard to interpret, SciPy's KS routine includes the functionality to convert the KS values into $p$-values, where the null hypothesis is that the test distribution and the distribution of the data are the same. $p$-values of greater than 0.05 mean that this null hypothesis cannot be rejected. Therefore, high $p$-values here loosely mean that the data is likely to come from the test distribution. This meaning of the $p$-value may seem somewhat backwards compared to usually adopted interpretations, and is due to the slightly unusual choice of null-hypothesis in the SciPy KS software. Nevertheless, high $p$-values here indicate the gamma distribution is a likely match for the data. All $p$-values in figure 5.3 are $> 0.2$, indicating, under this framework, a good fit to the gamma distribution.

These results demonstrate that across a broad spectrum of random networks, the resilience $\rho$ is log-normally distributed across network realisations. This finding will come into play in chapter 7, where it will be used to provide a computationally fast method of simulating cascades.

### 5.3.2 Causes of edge removal

5.3.1 showed that the mean fraction $\mathcal{S}$ of edges surviving a cascade varies sigmoidally with edge capacity $\alpha$. In this section, we look at what happens to the edges that do not survive. Following the cascade algorithm in figure 5.1, edges may be removed from the network either because their flow exceeds their capacity $\alpha$, or because the adjoining nodes have de-synchronised and drifted too far from the grid's reference frequency. Figure 5.4 shows the fraction of edges failing via each of these two modes for various network configurations as a function of edge capacity. The figure shows that at the critical point $\alpha_c$, networks whose compositions are indicative of a traditional electrical grid, result in the proportion of edges removed due to de-synchronisation exceeding those removed due to overload. The opposite is true for grids with equal numbers of sources and sinks. This indicates that networks with equal numbers of sources and sinks are less prone to de-synchronisation. This agrees with the findings in chapter 4, which showed that increased distribution results in a decreased critical coupling strength $\kappa_c$.

### 5.3.3 Duration of cascades

We now turn our attention to the question of how long cascades take to unfold in networks. The steady state model used in chapter 3, had no notion of time; it was just an iteration of steady states. The dynamical cascades used in this chapter, however, depend upon the time variable $t$, which has units of seconds. This allows the duration of a cascade to be measured. An understanding of the timescale of cascades will allow us to determine the possible impact of transients. If the cascade timescales are similar to those of the transient dynamics identified in chapter 4, then this indicates

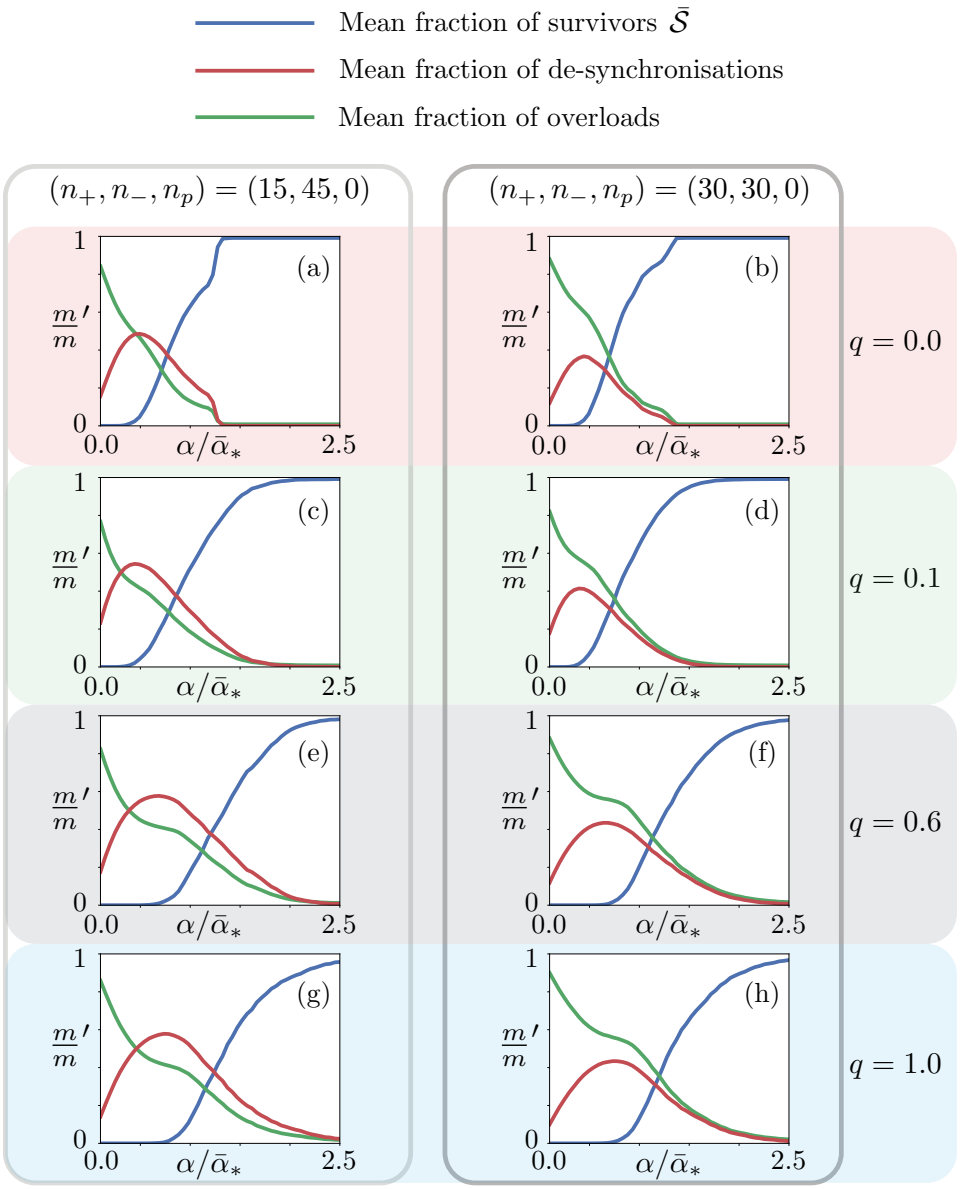



Figure 5.4: Fraction of edges, denoted $m'/m$, that either survive, overload, or de-synchronise, for 8 different ensembles, each containing 500 Watts-Strogatz networks. All axes have been scaled by the mean operating capacity $\bar{\alpha}_*$. The left-hand column of distributions is for ensembles with a node configuration characteristic of traditional grids with a few large generators. The right-hand column is for networks with equal numbers of consumers and generators, indicative of a modern distributed grid. Each of the four rows of panels represents a different value of randomness $q$, from regular lattices at the top to Poisson networks at the bottom.

that transients may play an important role in the evolution of a cascade, ultimately causing a divergence from the steady state cases of chapter 3. Understanding the time duration of cascades will also give intuition about the type of counter-measures available to protect against an ongoing cascade. If the durations are very short, then a network controller will have little time to respond and counter-measures must therefore be preventative rather than reactive. If the durations are very long, then a network controller may in principle be able to act to arrest the cascade.

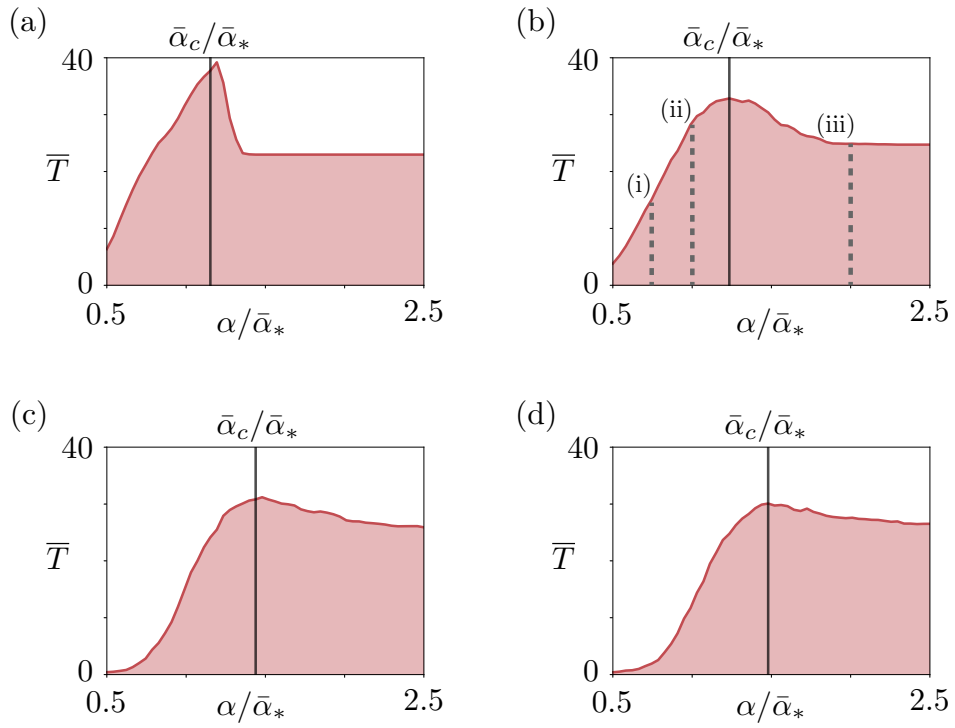



Figure 5.5: The mean duration $\overline{T}$, in seconds, of cascading failures as a function of the network edge capacity $\alpha$, normalised by each ensemble's mean operating capacity $\bar{\alpha}_*$. Each plot is for an ensemble of 500 Watts–Strogatz networks, each with $n = 50$ and $(n_+, n_-, n_p) = (10, 40, 0)$. The values of topological randomness $q$ in panels (a) to (d) are, respectively, 0, 0.1, 0.6 and 1.0. The solid vertical line in each plot is the location of the mean critical capacity. The underlying distributions at each of the vertical lines in panel (b) are plotted separately in figure 5.6.

Let us define the duration $T$ of a network cascade as the time taken in seconds from the initial removal of the most loaded edge, to the time at which the network reaches a steady state satisfying equation (5.7). Figure 5.5 shows the

mean time taken $\overline{T}$ for a cascade to occur, measured as a function of edge capacity, for four different network ensembles. In all the examples of figure 5.5, the duration of the cascades is short for low values of $\alpha$. From figure 5.2, we know that cascades in this region of $\alpha$ will generally be catastrophic, corresponding to values of $\bar{\mathcal{S}} \approx 0$. As $\alpha$ increases, which is associated with increasing $\bar{\mathcal{S}}$, the durations also increase. The duration reaches a peak at approximately the mean critical capacity $\bar{\alpha}_c$. Beyond this point, the duration settles to a value of around $\overline{T} \approx 30$s. Cascades therefore last the longest when approximately half the network survives, which occurs at $\bar{\alpha}_c$. Catastrophic network cascades, in which the entire network is destroyed, are the fastest, lasting only a few seconds.

Short durations therefore tend to occur for values of $\alpha$ associated with small values of $\overline{\mathcal{S}}$. Longer durations would seem to coincide with values of $\alpha$ that allow for greater survivability. The ensemble average durations $\overline{T}$ describe quite smooth profiles as a function of $\alpha$, but, as with $\rho$, individual values of $T$ within an ensemble are realisations of a random variable.

These features can can be seen more clearly by plotting individual values of $\mathcal{S}$ against $T$, for specific values of $\alpha$. Such scatter plots are shown in figure 5.6. Each of the four panels shows $\mathcal{S}$ plotted against $T$ for four different values of $\alpha$ in the same ensemble. This ensemble is in fact the same as that in figure 5.5(b). Each panel in figure 5.6 corresponds to one of the four vertical lines in figure 5.5(b), and reveals the underlying structure. Panel 5.6(a) shows $\mathcal{S}$ versus $T$ for $\alpha = 0.75$, which is beneath the critical value $\alpha_c$ and corresponds to a situation where most cascades result in a low value of $\bar{\mathcal{S}}$ and $\overline{T}$. This is confirmed in 5.6(a), which shows most of the cascades are tightly clustered in the bottom-left of the plot. These cascades completely destroy the networks, giving values of $\mathcal{S} \approx 0$ and very short durations $T$. However, it is also clear that there is another cluster of less severe cascades, whose durations are longer. The result is that the durations are in fact bi-modal, as shown by the adjoining histogram in 5.6(a). This same trend is also clear in 5.6(b), which shows the case for $\alpha = 1$. This is still beneath the critical value. The second cluster has now grown and encompasses most of the cascades. Panel 5.6(c) shows the situation at the critical point $\bar{\alpha}_c$. Here, a third cluster has emerged

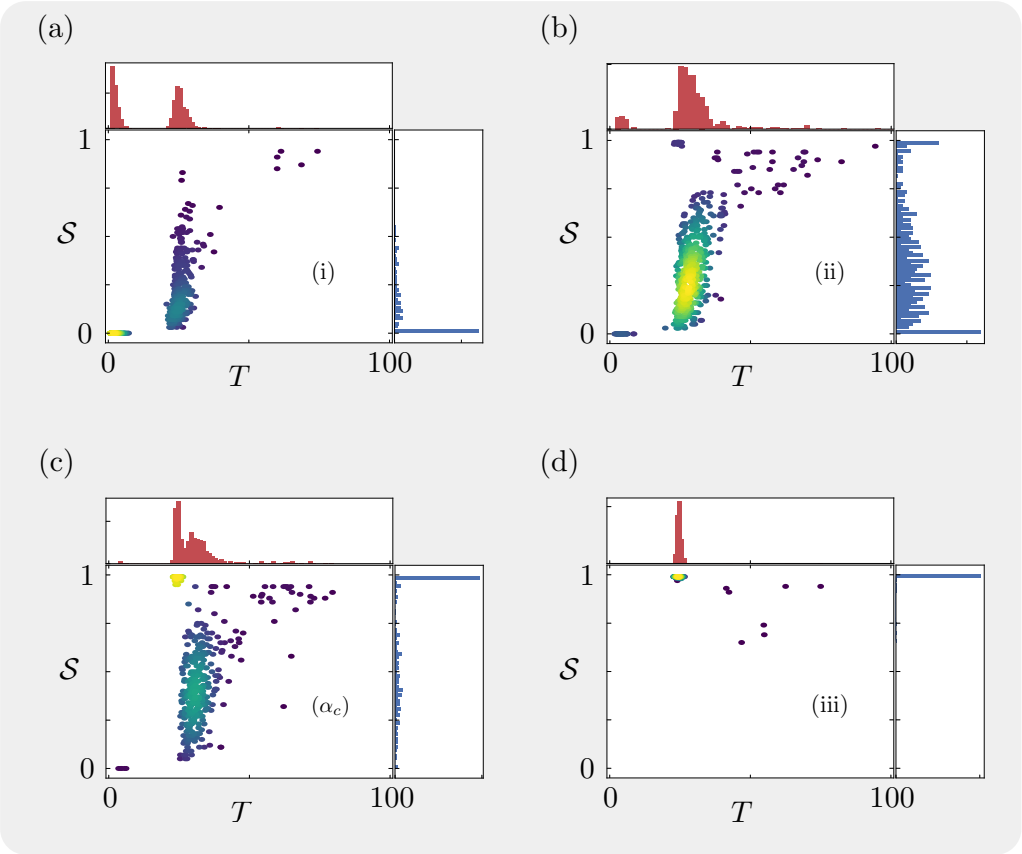



Figure 5.6: Scatter plots of the duration $T$ of a cascade against the fraction $\mathcal{S}$ of surviving edges, for an ensemble of 500 random networks with configuration $(n_+, n_-, n_p) = (10, 40, 0)$. This is the same ensemble as shown in figure 5.5(b). Each panel has a different value of edge capacity $\alpha$, corresponding the vertical sections in 5.5(b). The values of $\alpha$ for panels (a) to (d) are, respectively, 0.75, 1, 1.21, and 2. The colours of the scatter points encode the density of points; the brighter the denser. Adjoining each axis are histograms giving a picture of the distribution of each variable.

at the top centre-left which now contains the vast majority of the cases. This can be seen as the very small, bright yellow blob. These are the cascades which cause little damage, and whose durations are roughly the same as the second cluster of intermediately damaging cascades. For $\alpha$ greater than $\bar{\alpha}_c$, as shown in 5.6(d), the only cluster of any significance remaining is that of the cascades causing little damage, which emerged in panel (c).

In summary, a short duration is associated with catastrophic cascades, while cascades that cause little impact have a longer and approximately constant

duration. The duration is maximised at the critical value $\bar{\alpha}_c$. Additionally, the values of $\overline{T}$ mean that cascades unfold on a similar timescale to the transients identified in chapter 4, further motivating the use of a dynamical model for determining resilience.

### 5.3.4 Effect of increasing connectivity $K$ on $\rho$

In chapter 3, it was found that for steady-state networks containing only one large source, and many small sinks, the resilience increased rapidly with increased degree $K$. This was indicated by the rapidly decreasing value of $\bar{\rho}$. It was further identified that this decrease in $\bar{\rho}$ could be well explained by the function $K/(K-1)$. To examine if this behaviour holds when considering

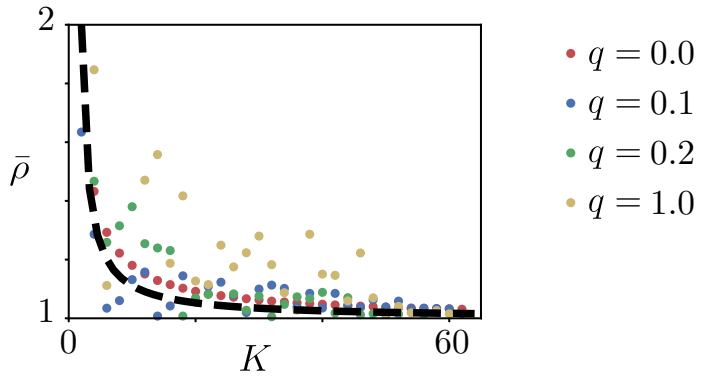



Figure 5.7: Mean resilience $\bar{\rho}$ as a function of mean degree $K$ for networks all with configuration $(n_+, n_-, n_p) = (1, 59, 0)$. Each data point is obtained from an ensemble average across 200 network realisations. The dashed line indicates the function $K/(K-1)$.

dynamical networks, $\bar{\rho}$ has been plotted as a function of $K$ in figure 5.7 for various ensembles of networks each containing only one large source. It can be seen that the $\bar{\rho}$ values do still decrease with $K$. However, the relationship to $K/(K-1)$ is no longer as clear, particularity for the more random networks with $q = 1$. The data is significantly noisier than the steady state equivalent. The $K/(K-1)$ approximation does however remain a satisfactory guideline for the behaviour of regular lattices ($q = 0$), with only one large source.

These results mean that while increasing $K$ does still increase resilience, it is a less efficacious strategy than in the steady state case; this is especially true for more random networks.

## 5.4 Dependence of $\rho$ on node composition

The dependence of resilience upon the proportions of consumers and generators will now be investigated by computing $\bar{\rho}$ across ensembles of random networks. Following the methodology of previous chapters, $\bar{\rho}$ will be projected onto the node configuration simplex. Recall that this simplex is formed by the conditions $n_+ + n_- + n_p = n$, $n_+ \geq 1$, and $n_- \geq 1$. Each point on the simplex represents one of the $n(n-1)/2$ possible node configurations $(n_+, n_-, n_p)$ for networks of fixed size $n$. For a schematic of this simplex refer back to figure 3.10 in chapter 3. These configuration simplexes provide a neat visualisation of the variation in resilience with the numbers and ratios of source and sink nodes. Regions of the simplex with low values of $\bar{\rho}$ indicate node configurations $(n_+, n_-, n_p)$ offering high resilience.

Figure 5.8 shows simplex plots for ensembles of random networks generated using the Watts-Strogatz method (Watts and Strogatz, 1998). These networks are parameterised by their size $n$, their mean degree $K$, and their rewiring parameter $q \in [0, 1]$ which interpolates between regular lattices at $q = 0$ and Poissonian networks at $q = 1$. All the simplexes shown in figure 5.8 have $n = 100$ and $K = 4$, with each pixel representing the value of $\bar{\rho}$ obtained from an ensemble of 100 network realisations. All of the simplexes appear somewhat noisy, especially compared to their steady state counterparts obtained in chapter 3. This can be attributed to the smaller ensemble sizes used to generate the simplexes in figure 5.8. These smaller ensemble sizes were required due to the increased computational workload involved in computing a single value of $\bar{\rho}$ in these dynamic networks. Nevertheless, clear trends can still be identified from the data in figure 5.8. The simplex in panel 5.8(a) shows the case for networks with $q = 0$. Here, the highest values of $\bar{\rho}$ are found at the apex of the simplex and down the two lateral edges. This means that networks with only one large generator (or equivalently one

large consumer, due to source-sink reciprocity) have the least resilience. The highest resiliences, given by the lowest values of $\bar{\rho}$, are found in the interior of the simplex. This region represents networks with a highly distributed mix of small consumers and generators. The absolute lowest value is found at

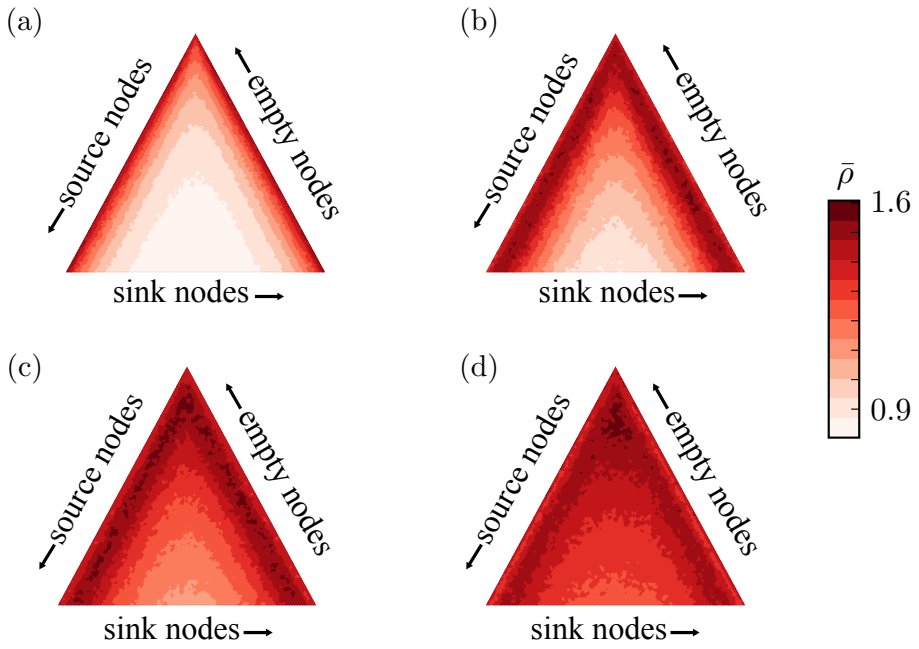



Figure 5.8: Panels (a) to (d) show $\bar{\rho}$ projected onto the node configuration simplex for Watts-Strogatz networks with topological randomness parameter $q = 0$, 0.2, 0.4 and 1 respectively. Simplexes (a) to (d) therefore represent increasing structural randomness from regular lattices in (a) to Poisson networks in (d). All networks have $n = 100$ and $K = 4$. Each pixel shows $\bar{\rho}$ obtained from an ensemble average over 100 network realisations where source and sink locations are allocated randomly. $\bar{\rho}$ is computed using the algorithm outlined in figure 5.1.

the bottom-centre of the simplex, where networks possess equal numbers of generators and consumers with no empty nodes. This region of the simplex possesses values of $\bar{\rho} < 1$, meaning that the critical amount of edge capacity $\alpha_c$ required to survive a cascade is less than the operating capacity $\alpha_*$. The networks occupying this region are therefore inherently resilient to cascades and require no additional capacity. In totality figure 5.8(a) demonstrates that for regular networks, increased distribution of generation leads to substantial increases in resilience compared to networks with only a small number of large generators. The same trend can also be seen in figure 5.8(b) which

shows the case for networks with $q = 0.2$. These networks are therefore more topologically random than those in 5.8(a). The relationship between increased distribution and increased resilience is still clear, with the lowest values of $\bar{\rho}$ occupying the central region of the simplex. There are however two key differences between the results in figure 5.8(b) and 5.8(a). Firstly, the values in the centre of simplex 5.8(b) are higher than their counterparts in 5.8(a), having values of $\bar{\rho} \approx 1$. Secondly, the region of low resilience extends further into the interior of the simplex compared to 5.8(a). Taken together these two observations mean that for networks with $q = 0.2$, higher numbers of small sources and sinks are required to achieve substantial increases in resilience and that ultimately these increases are less than in regular lattices. As the networks become progressively disordered, shown by the cases for $q = 0.4$ in figure 5.8(c) and $q = 1$ in 5.8(d), the benefits of distribution continue to decrease. For the completely disordered Poisson networks of 5.8(d), the values of $\bar{\rho}$ in the interior of the simplex are almost homogeneous with values of $\bar{\rho} > 1$ indicating that networks with equal numbers of consumers and generators are no longer inherently resilient. Some of the highest values of $\bar{\rho}$ can be found in the upper-centre of the simplex, indicative of networks with a small but equal number of consumers and generators, and a large number of empty nodes.

In summary, the simplexes in figure 5.8 show that for more ordered networks, an increased number of small, distributed generators gives a substantial boost in the relative resilience as gauged by $\bar{\rho}$. However, as networks become more random and disordered, increasing the level of distribution provides less benefit. Equivalent plots for networks of size $n = 50$ are shown in figure 5.9, demonstrating that these results also apply to smaller networks.

The results in figure 5.8 can be directly compared to the simplexes obtained previously in chapter 3, figure 3.10, which were computed using a steady state flow model. Those simplexes also identified that increased distribution delivers increased resilience for ordered networks, but that as disorder increases the benefits diminish. However, this trend was much more pronounced in the steady state case compared to the dynamic case of figure 5.8. Additionally, the morphology of the simplexes in figure 5.8 also differs slightly from the

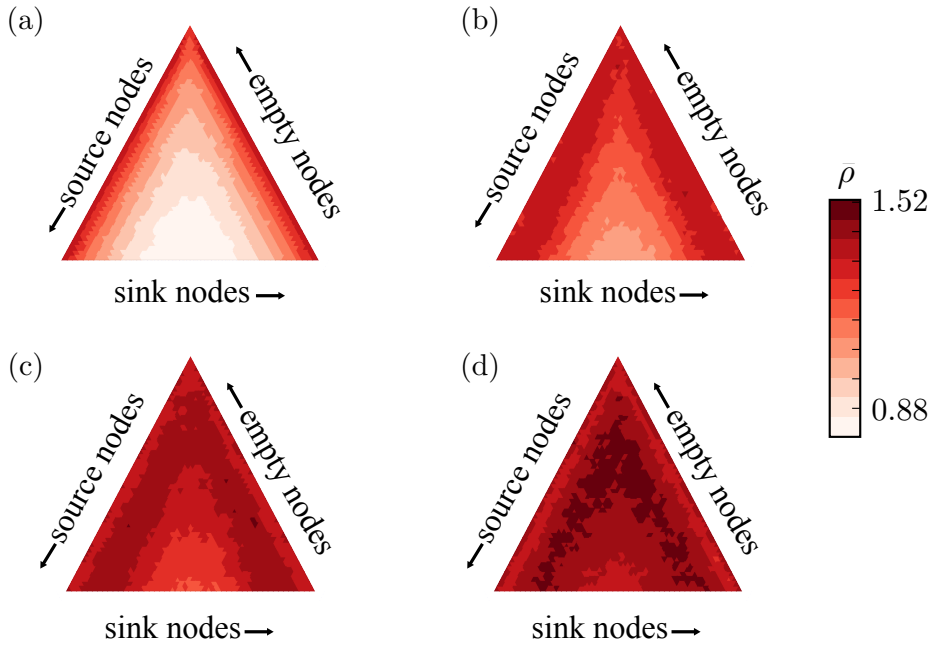



Figure 5.9: Panels (a) to (d) show $\bar{\rho}$ projected onto the node configuration simplex for Watts-Strogatz networks with topological randomness parameter $q = 0$, 0.2, 0.4 and 1 respectively. Simplexes (a) to (d) therefore represent increasing structural randomness from regular lattices in (a) to Poisson networks in (d). All networks have $n = 50$ and $K = 4$. Each pixel shows $\bar{\rho}$ obtained from an ensemble average over 100 network realisations where source and sink locations are allocated randomly. $\bar{\rho}$ is computed using the algorithm outlined in figure 5.1.

steady state case. In particular, the most resilient values are found at the bottom of the simplexes, indicative of networks with no passive nodes. By contrast, in the steady state case the lowest values were found in the upper centre of the simplex, indicating that having some amount of passive nodes conferred additional resilience. Therefore, in the dynamic case, while increased distribution still in general boosts resilience, passive nodes are no longer advantageous. The dynamic, transient properties of the swing equation therefore have a significant impact on the resilience of networks which must be taken into consideration when assessing the optimal operating composition of an electrical network.

## 5.5 Conclusion

This chapter has introduced a model of dynamical cascading failures. The basis of the model is the swing equation, and allows the dynamical properties of power flow to be captured and taken into account when assessing the resilience of a power grid to a cascade. This chapter has shown that for ensembles of random networks, the resilience measure $\rho$ used previously in chapter 3 is still a meaningful and measurable quantity. Additionally, $\rho$ is log-normally distributed across random networks; this convenient statistical property will be exploited in the next chapter.

The objective of using the the dynamical model in this chapter was to investigate if the transient properties of dynamic power flow significantly impact the resilience of networks, causing a divergence from the resilience results of chapter 3. By examining the duration of cascades in each network, this chapter has shown that the timescale of cascades does indeed coincide with those of the transients identified in chapter 4. This means that we should expect the dynamics to have a significant impact on cascades. It has also been identified that large proportions of edges also fail due to de-synchronisation, rather that capacity overload. Higher levels of distributed generation were found to reduce the amount of these de-synchronisations.

Examining the durations of cascades also identified that the more catastrophic the cascade, the shorter its duration. This offers a poor outlook for network operators who may wish to interrupt and arrest a cascade, since the most damaging events are typically over in only a few seconds. This adds extra importance to preventative measures; in particular, the design of grids that are inherently robust. Simply increasing the mean degree of networks, likely to be an expensive option, has been shown to only offer limited additional resilience. However, adjusting the proportion of generators and consumers in the networks has been shown to have a substantial ability to boost the resilience, particularly for more ordered network topologies. This is an encouraging finding for the design of modern distributed grids, whose numbers of generators and consumers can typical be varied and controlled (Olivares et al., 2014). The next chapter applies some of these findings to test cases

obtained from real-world generation and consumption data.

# Chapter 6

# Micro-grid test case

## 6.1 Introduction

The continuing shift towards renewable generation is causing an overhaul in the structure of power grids (D'Souza, 2017). At the macroscopic level, high power transmission grids are becoming more decentralised as wind and solar power farms come online. At the micro level, household scale distribution networks are increasingly variable and even passive as houses become partially self-sufficient through solar panels and batteries. The composition of these distribution grids may change throughout a day resulting in a highly dynamic grid with an extensive operating space.

The previous chapters have used both steady state and dynamical models to characterise the efficiency and resilience of flow networks. Chapters 3 and 5 showed that the resilience of networks to cascading failures varies as a function of the proportion of consumers and generators in the network. Chapter 4 demonstrated that the critical coupling strength required for grid synchronisation also varies with consumer-generator composition. This chapter applies these findings and their underlying methods to a data-driven model of an electrical micro-grid. Power consumption and generation data from UK households is used to show how variability in a real-world energy landscape impacts the resilience of the networks.

## 6.2 Data-driven micro-grid model

Power consumption and generation data will be used throughout this chapter to construct models of electrical micro-grids. These models will provide an experimental test bed in which to investigate how the inherent variability of renewable power impacts the resilience of the grid. This section introduces the principles of micro-grids, and then provides an overview of the relevant data sets and how they will be analysed using the methods of the previous chapters.

### 6.2.1 Overview of micro-grids

A micro-grid is a collection of houses, each of which may possess photo-voltaic (PV) panels and batteries. These houses are networked together with connections capable of supporting bidirectional power flow. As the name suggests, micro-grids are typically small and consist of some tens of households. These small bidirectional power networks are designed to support self-sufficiency; households may be able to use their own PV production to match their consumption needs, or even supply excess production back to the grid. Each micro-grid is connected to the external regional distribution grid via a node in the micro-grid called a principle point of connection (PPC). The PPC acts to balance demand within the network. If there is excess power being generated within the micro-grid, then the PPC functions as a sink node and consumes the excess production. If demand in the micro-grid is exceeding local PV generation, then the PPC functions as a source node and supplies power into the micro-grid from the external distribution network. If net demand on the PPC is zero, then the grid is self-sufficient. This situation is often referred to as *islanding*. For a summary of micro-grid designs and their development, refer to Parhizi et al. (2015). Micro-grids are widely regarded as an efficient and resilient design strategy for future low-carbon grids and are already being implemented in sustainable housing projects. For an example of how conventional grids might be converted into ensembles of micro-grids, refer to Halu et al. (2016).

### 6.2.2 Overview of data

In order to construct synthetic models of micro-grids, we require both real-world household power consumption data as well as photo-voltaic (PV) power output data. Together, these will give a realistic model of the daily variability in power flow. For power generation, we shall be using the data set curated by UK power networks and London city council. This data set is available from London Datastore (2014). It was gathered over the course of approximately two years and comprises the power consumption levels of around 5000 houses in the greater London area. The data was collected from smart meters installed in user's houses, and provides half-hourly resolution. For PV generation data we use another data set also curated by UK power networks, available at London Datastore (2016). The data contains the power output for an ensemble of low-voltage domestic scale PV units, also in the greater London area. This PV data contains over a year's worth of measurements at up to 10-minute resolution windows.

Before detailing how these two data sets will be used to generate synthetic micro-grids, let us first take an overview of the data. Figure 6.1 shows the mean power consumption of all households in the power consumption data set. In particular, panel 6.1(a) shows the mean weekly power consumption measured in kWh per week over the entire range of the data, from January 2012 to January 2014. Clear seasonal trends are visible, with the highest power demand occurring in the winter of each year. This trend is common for the UK, where winters have long nights and central heating demand is high. In regions with hot summers and high usage of air conditioning, such as the southern US, the opposite trend may be true with highest consumption occurring in the summer. Panel 6.1(b) focuses on the month of December 2013, as highlighted by the grey shaded region in panel (a). This month-long view of power consumption in kW shows the regular daily oscillations in demand. Panel 6.1(c) focuses on the week of 07/12/13 to 14/12/13, as highlighted by the grey region in panel (b), and shows these daily oscillations in higher detail. The morphology of these oscillations appear regular. The minimum consumption each day occurs in the middle of the night at around 3am. There is then an increase in the morning, which sometimes manifests

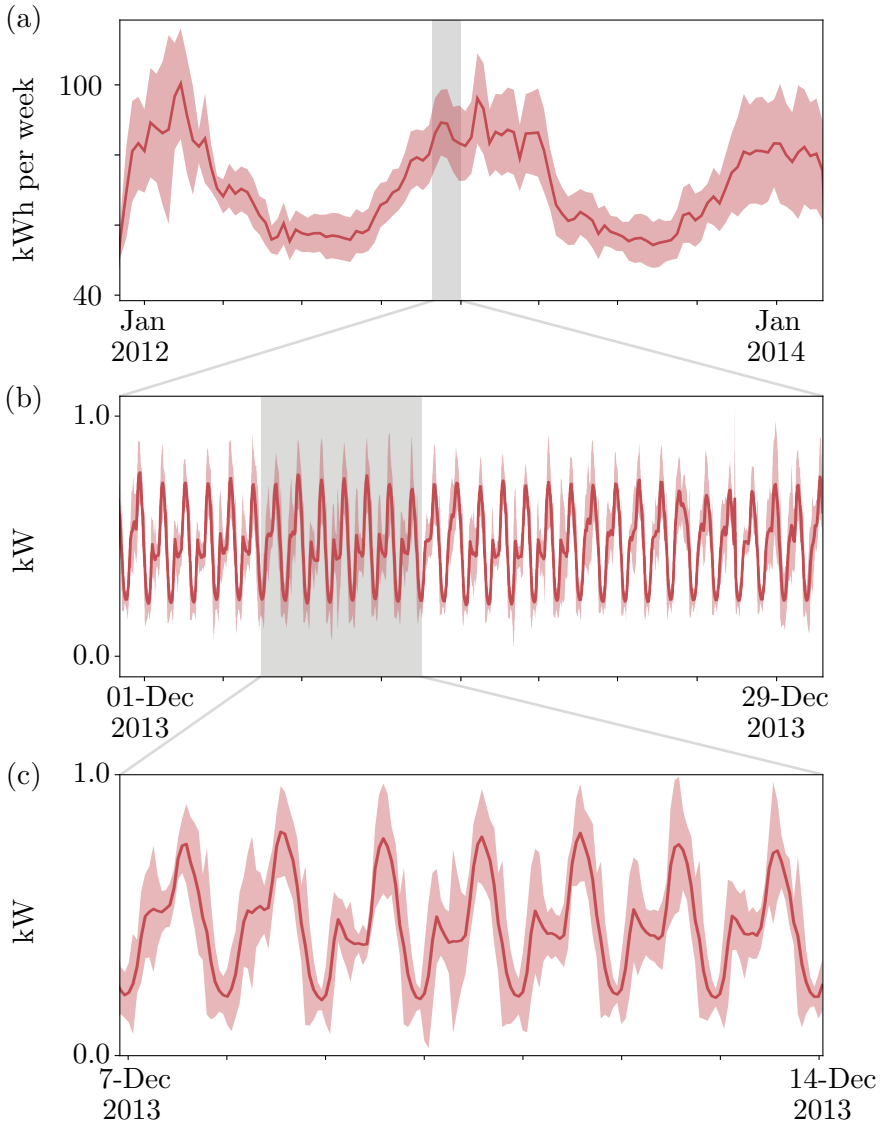



Figure 6.1: Mean power demand over all houses in the data set from (London Datastore, 2014). Red shaded regions indicate standard deviations. Panel (a) shows mean total weekly power demand measured in kWh per week. Panel (b) shows power consumption in kW over the month of December 2013 and (c) shows power consumption over the second week of that month. The first two days are the weekend. The ticks along the horizontal axis in panel (c) denote 3am each day, which approximately corresponds to the daily nadir in power consumption. Each panel is a zoomed in version of the panel above.

as a secondary peak in consumption. This may correspond to people people waking up and the activation of timed central heating systems. The demand then dips around midday when most people are out at work, before reaching

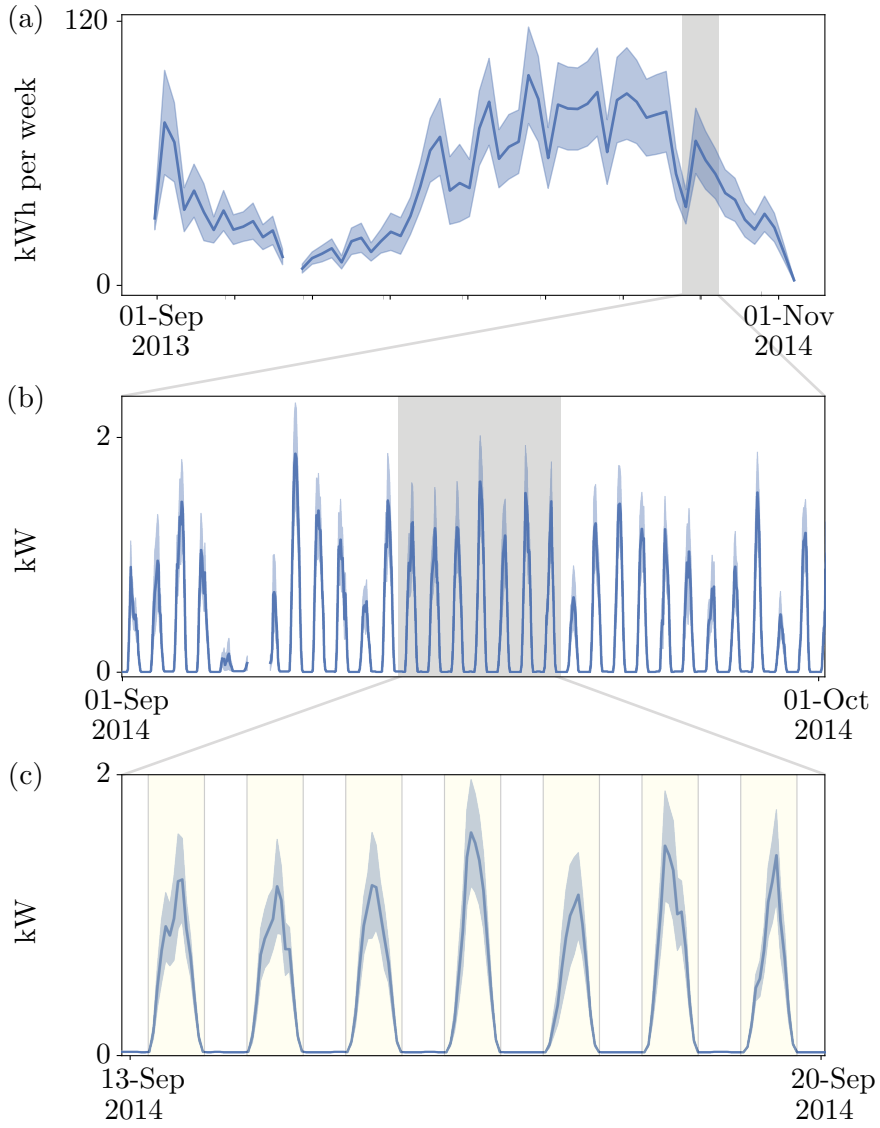



Figure 6.2: Mean power generation over all PV panels in the data set from (London Datastore, 2016). Blue shaded regions indicate standard deviations. Panel (a) shows mean total weekly power generation measured in kWh per week. Panel (b) shows power generation in kW over the month of September 2014 and (c) shows power generation over a week in that month. Light shaded yellow areas in (c) denote hours of sunlight (as obtained from timeanddate.com, 2021). Each panel is a zoomed in version of the panel above.

a maximum peak in the evening when families return home.

Figure 6.2 shows the mean power output of all PV panels in the generation data set. Panel 6.2(a) shows the weekly mean power generation measured in

kWh per week from September 2013 to November 2014. Unsurprisingly the peak power generation occurs during the summer months, and is minimal during the darker winter months. Panel 6.2(b) focuses on the month of September 2014, as highlighted by the grey region in panel (a), and clearly shows the daily oscillations in power output. The amplitudes of these oscillations are less regular than those in the power consumption data, highlighting the impact of weather variability. Panel 6.2(c) focuses on the week of 13/09/14 to 20/09/14, as highlighted by the grey region in panel (b), and demonstrates morphology of each day's power generation profile. Maximum generation occurs around midday, with no production occurring at night.

Taken together, these two data sets provide a realistic picture of the daily and seasonal variations in power usage in low voltage domestic settings. This will be utilised throughout the rest of this chapter to provide a realistic setting in which to deploy some of the analysis of previous chapters.

### 6.2.3 Data usage in micro-grid models

The models of micro-grids in this chapter comprise $n-1$ houses together with 1 PPC node, giving $n$ nodes in total. For each of the $n-1$ houses, a time-series of power consumption will be selected uniformly at random from the approximately 5000 time-series contained within the data set. A subset of the houses will be chosen to have PV panels. The fraction of houses with PV will be referred to as the PV uptake. A PV uptake of 100% means that all $n-1$ houses are equipped with PV panels. The output of each PV unit will also be given by a time-series selected at random from the data set. At each point in time, the power consumption and generation will be balanced via the PPC. Each such micro-grid is therefore a random realisation that displays realistic daily and seasonal fluctuations in power usage.

It should be noted that the time period covered by the power consumption data is not the same as the period covered by the PV data, although the two do overlap. However, this chapter is not concerned with reconstructing a particular historical period, nor with attempting to infer the impact of PV generation on the usage of power. All that is needed for this chapter is a

realistic notion of the daily variability of PV and household power usage, for which these two data sets are perfectly adequate. All time-series data for PV panels will be sampled from within the range November 2013 to November 2014 in the PV data set. All time-series data for household power consumption will be sampled from within the range January 2013 to January 2014 in the power demand data set. This gives a full years worth of data for both PV and consumption. For a typical micro-grid experiment in this chapter, its activity will be tracked over the course of a given week in this synthetic year. To constrain the focus of these experiments, four seasonal time windows will be considered: Winter, Spring, Summer, and Autumn. Winter will be represented by the first week in January, spring by the first week in May, summer by the first week in July, and autumn by the first week in October. This will allow the analysis of both the daily and seasonal variability in micro-grid activity.

### 6.2.4 Simplex projection

Recall that previous chapters used the node configuration simplex to examine the impact of varying the proportions of source and sink nodes in networks. Networks with $n$ nodes were constructed wherein $n_+$ nodes were sources having a positive power output, $n_-$ were sinks, and the remaining $n_p$ were passive. The condition $n = n_+ + n_- + n_p$ then allowed each network with composition $(n_+, n_-, n_p)$ to be represented as a unique point on the simplex. Refer back to figure 3.10(a) for a diagram of this configuration space. Projecting results onto the simplex allowed quantities like efficiency, critical coupling, and resilience to viewed as a function of node composition. This revealed the effect of altering node composition on these key characteristics of network function.

However, the effective node compositions of the micro-grids considered in 6.2.3 are not static. The constantly varying levels of household demand and generation result in a mutable node composition. On a winter's evening, there will be high power demand due to central heating and no PV output. This will result in most nodes being net consumers. Midday in the summer,

when there is little power demand and PV units are at their most productive, will yield a network of net generators. Additionally, micro-grids are likely to be relatively passive due to the heterogeneity of power demand amongst households. For instance, on a given evening there are likely to be a few households whose power demand significantly exceeds most others. There will also likely be a subset of houses that on a particular evening have little to no demand, perhaps because the occupants are absent. These features would result in a network where most of the demand is dominated by a subset of houses; other houses with much lower demand would act mostly as passive junctions.

In order to investigate the impact of this variable composition on network performance, let us now examine a method of projecting the state of a network onto the node configuration simplex. We wish to be able to express the composition of a network via three numbers, as in previous chapters where networks where characterised by $(n_+, n_-, n_p)$. First of all, let the power demand state of a micro-grid at time $t$ be given by the power vector $P \in \mathbb{R}^n$. If a node $i$ has power $P_i < 0$, then it is a net consumer. If it has $P_i > 0$, then it is producing more power via PV than it is consuming and is thus a net producer. Let us now define a source-node density $\eta_+$ as

$$\eta_+ := \frac{1}{n \max(P)} \sum_{x \in P^+} x, \tag{6.1}$$

where $P^+$ is the vector containing only the positive entries of $P$. Equation (6.1) therefore takes a sum of the source terms in the network and normalises it by the largest term. Equivalently let us define sink-node density $\eta_-$ as

$$\eta_- := \frac{1}{n \min(P)} \sum_{x \in P^-} x, \tag{6.2}$$

where $P^-$ is the vector containing only the negative entries of $P$. The passivity $\eta_p$ of the network is then defined as

$$\eta_p := 1 - \eta_+ - \eta_-. \tag{6.3}$$

This procedure therefore gives three numbers $(\eta_+, \eta_-, \eta_p)$ between 0 and 1

quantifying the density of generators and consumers in a network. These numbers also define a unique position on a simplex. The results displayed in simplexes in previous chapters are mostly relative quantities, such as resilience $\rho$, which have been normalised by power flow volume. Therefore those results can be used in conjunction with the $(\eta_+, \eta_-, \eta_p)$ mapping to examine the varying resilience of micro-grids.

Let us now examine some simple examples of networks comprising only 4 nodes in order to provide intuition about the $(\eta_+, \eta_-, \eta_p)$ coordinate system. First consider a 4 node network with power vector $P = (3, -1, -1, -1)^T$. There is one generator providing all the power to three smaller consumers. This is a composition representative of a traditional power grid and gives $(\eta_+, \eta_-, \eta_p) = (\frac{1}{4}, \frac{3}{4}, 0)$. The network therefore has no passivity $\eta_p = 0$, but a high sink density as the network is composed mostly of small sinks. Now consider the case with $P = (1, 1, -1, -1)^T$ where there are equal numbers of sources and sinks with and no passive nodes. This is represented as $(\eta_+, \eta_-, \eta_p) = (\frac{1}{2}, \frac{1}{2}, 0)$. Unsurprisingly, the source and sink densities are equal and there is no passivity. As another example, consider the network with $P = (1, 0, 0, -1)^T$ containing one source, one sink, and two empty nodes. Correspondingly we have a simplex coordinate $(\eta_+, \eta_-, \eta_p) = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$ containing a dominant passive term. As mentioned earlier, sink heterogeneity can lead to high passivity. To see this consider the example of a network with power vector $P = (12, -9, -2, -1)^T$. Here there is one large source with output $+12$ and three sinks. The demands of the sinks are highly skewed, with one large sink having a power demand of 9 and the two others being less than a third in strength. These small sink nodes are therefore relatively passive compared to the dominant sink. The corresponding coordinate is $(\eta_+, \eta_-, \eta_p) = (\frac{3}{12}, \frac{4}{12}, \frac{5}{12})$, which has a large passive term.

The composition of a micro-grid will change throughout the day as demand and production vary, tracing a trajectory through the $(\eta_+, \eta_-, \eta_p)$ simplex. To demonstrate this, figure 6.3 shows the trajectory of a day in September for a realisation of a micro-grid derived from the data. It has 25 nodes and all houses are equipped with PV. Midnight is represented by point (i) in the simplex in figure 6.3(a). At this point, there is no PV production but

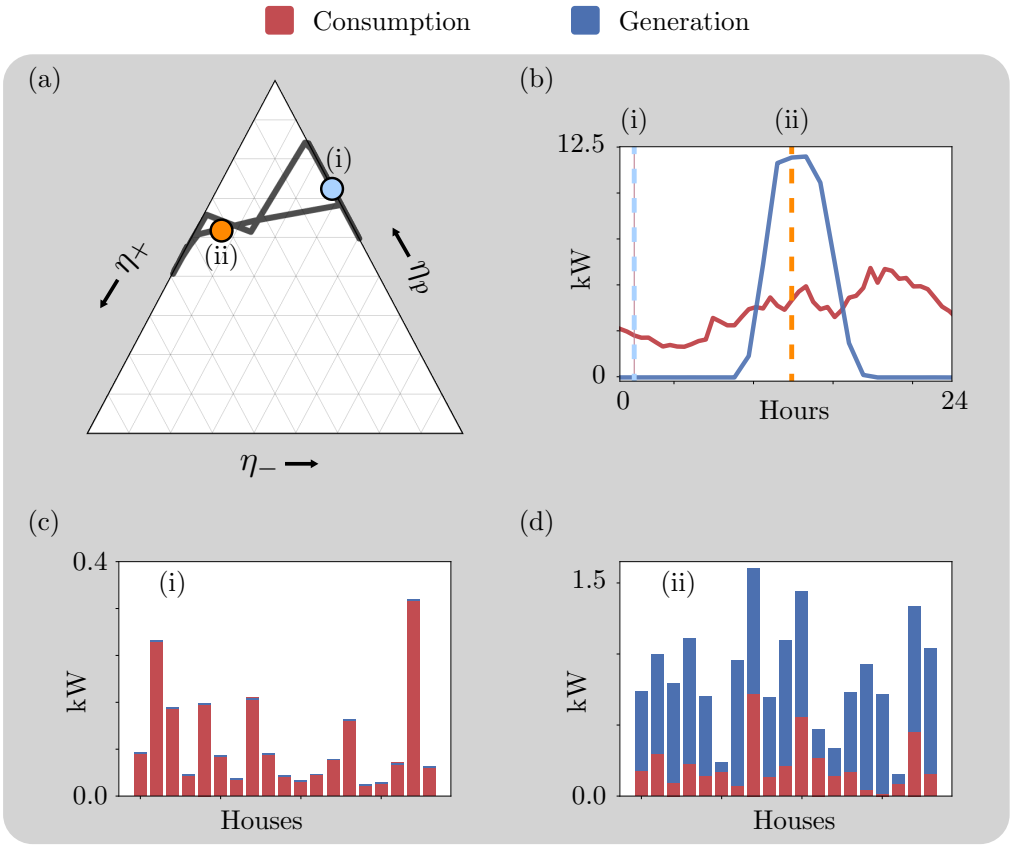



Figure 6.3: The daily trajectory in the node composition simplex of a microgrid with $n = 25$ nodes, where all houses have PV. The data for each house is sampled from September, giving a trajectory representing late summer in the UK. Panel (a) shows the trajectory in the simplex whose axes are defined in equations (6.1) to (6.3). Points (i) and (ii) represent midnight and midday respectively. Panel (b) shows the total consumption and generation in the network over the 24 hours, with the positions of (i) and (ii) indicated by the dashed lines. Panels (c) and (d) show the house-by-house power activity at midnight and midday respectively.

still a significant amount of demand, as shown by the demand-production plots in figure 6.3(b). The demand at midnight is highly heterogeneous, as shown by the household power activity chart in panel (c). This heterogeneity manifests as high passivity; correspondingly, point (i) is located high up in simplex (a) in the region indicative of highly passive networks. The location of point (i) on the right edge of the simplex indicates that the network is dominated by sink terms; the only source is the PPC which must supply all of the power since there is no PV online. Midday is indicated by point (ii) in

simplex (a). At this point there is a high volume of PV production which is outstripping the demand, as shown in panel (b). Again, the household activity chart in panel (d) shows that the demand and production is highly heterogeneous resulting in point (ii) being located high up in the simplex. Point (ii) is located towards the left edge of simplex (a), indicative of a network dominated by source terms. This is indeed the case at midday, where most houses are net sources. The daily trajectory therefore starts on the right edge, where there is no generation and middling demand, before swinging across to the opposite site of the simplex towards midday where the network is mostly generative. The trajectory then swings back over to the right hand side at night.

The daily oscillation in the simplex, from the right hand side at night to the left hand side in the day, appears to be a general behaviour of micro-grids. Further examples are demonstrated in figure 6.4. Each panel in the figure represents the ensemble average weekly trajectory of 50 micro-grid realisations, each constructed using the data. The left column of panels represents networks where only half of the houses are equipped with PV, while the right hand column represents networks with 100% PV uptake. Each row represents a week from each season of the year. In each case, the trajectories spend the nights on the upper right edge, before departing into the centre of the simplex in the morning and travelling towards the left edge. This movement becomes more pronounced as the seasons progress and daylight hours increase, leading to greater PV output. For example, panel (f) shows that in the summer the full uptake networks travel over halfway down the left edge of the simplex indicating that generation is significantly outstripping demand. The characteristic oscillations are more pronounced in the full PV uptake panel. The smallest movement in the simplex is observed in the winter for the half uptake case in panel (a), where the trajectories barely depart from the upper right corner. This indicates a highly heterogeneous demand with little to no production. In all cases, the trajectories live mostly in the upper half of the simplex corresponding to high passivity and therefore high heterogeneity. This passivity is typical of micro-grids. Performing similar simplex projections on macroscopic transmission grids, where all nodes represent large, similarly rated substations would likely

produce trajectories living much lower down in the simplex indicative of their relative homogeneity.

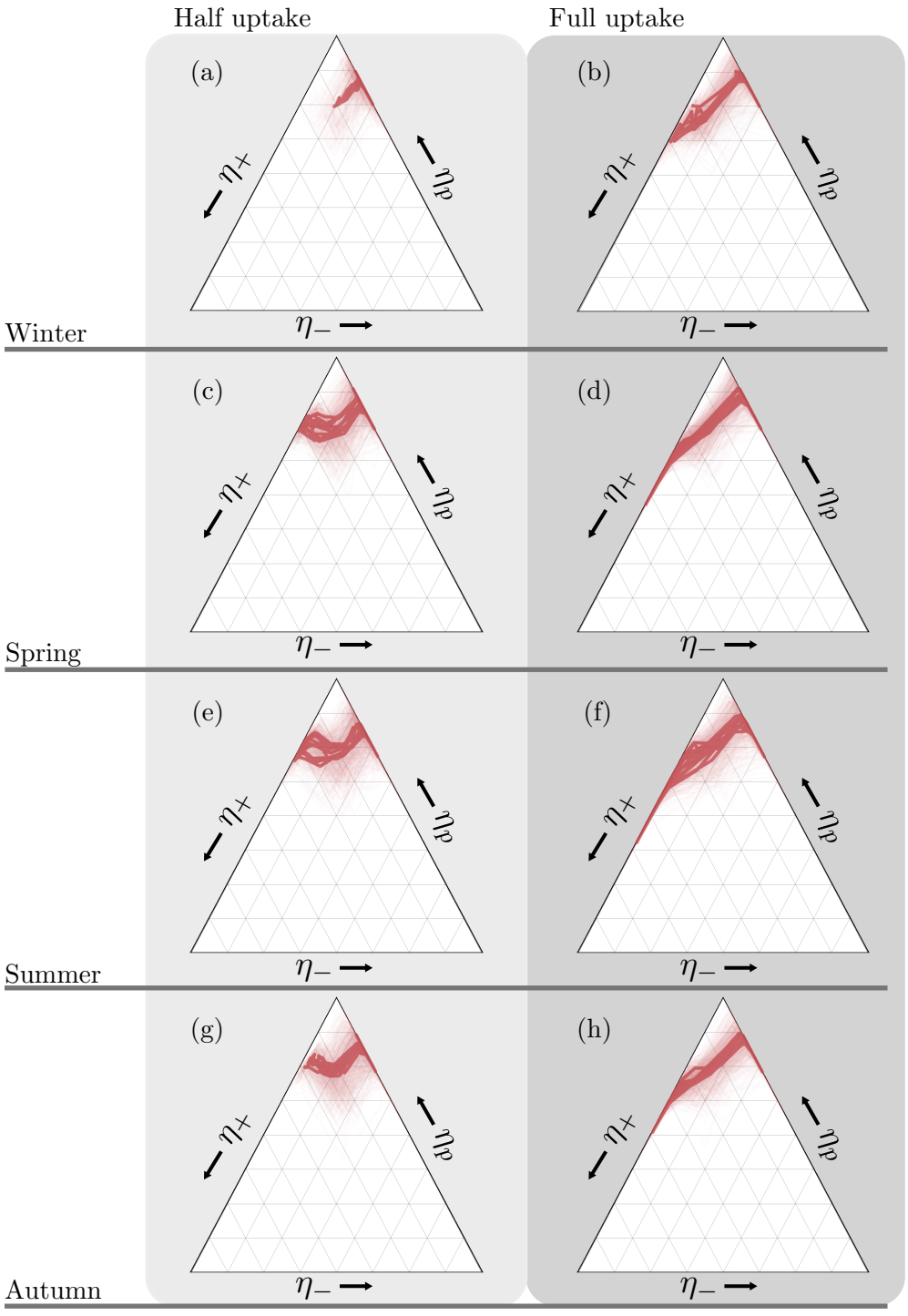



Figure 6.4: The thick red lines indicate the mean weekly trajectories for ensembles of 50 micro-grids with $n = 50$ nodes, each constructed using the power consumption and demand data-sets. The left hand column represents networks where only half of the houses have PV panels, and the right where all houses have PV. Each row represents a different season.

## 6.3 Variability in critical coupling

Power grids must operate within a narrow frequency range. Any significant drift from the grid's reference frequency will result in physical damage to grid components. Maintaining grid synchrony is therefore vital for the safe operation of power infrastructure. Chapters 4 and 5 modelled power grids using the swing equation, a model capable of capturing frequency dynamics. The critical coupling strength $\kappa_c$ is the minimum coupling necessary to achieve grid synchrony in the swing system. Low values of $\kappa_c$ are desirable; the lower its value, the more the network composition inherently facilitates synchronisation. For two and three node swing equation systems, the value of $\kappa_c$ can be found analytically (Manik et al., 2014). Chapter 4 used numerical simulations to find $\kappa_c$ in complex networks and showed that its mean value $\bar{\kappa}_c$ varies as a function of the composition of source and sink nodes in a network. As discussed in previous sections, micro-grids have a time-varying source-sink node composition. It should therefore be expected that the critical coupling will also vary with time. This section tracks the composition of ensembles of micro-grids over the course of a week, and uncovers these variations in $\bar{\kappa}_c$.

The variability of $\bar{\kappa}_c$ will be evaluated as follows. First, an ensemble of $\mathcal{N}$ micro-grid models will be instantiated. For each micro-grid, the effective node composition $(\eta_+, \eta_-, \eta_p)$ will then be computed at each time point via the procedure in section 6.2.4. Time points are at 10 minute intervals. Linear interpolation will be used to provide the necessary 10 minute resolution in each time-series. This will result in $\mathcal{N}$ trajectories through the $(\eta_+, \eta_-, \eta_p)$ simplex. These trajectories will then be averaged at each time point to give a single, ensemble mean trajectory $\mathcal{T}$. Each point in trajectory $\mathcal{T}$ will be a coordinate $(\eta_+, \eta_-, \eta_p)$ in the node composition simplex. We now wish to find the value of $\bar{\kappa}_c$ for each of these points. Recall that the simplexes in figure 4.3 in chapter 4 show $\bar{\kappa}_c$ for ensembles of networks of size $n = 50$, whose topologies are generated using the Watts–Strogatz procedure. Mapping the trajectory $\mathcal{T}$ onto these simplexes will reveal the variation in $\bar{\kappa}_c$. To do this, each point in $\mathcal{T}$ will be multiplied by $n = 50$, since $(\eta_+, \eta_-, \eta_p)$ coordinates are densities, as computed via equations (6.1), (6.2), and (6.3). This will yield the effective position of each trajectory point in the simplexes of figure

4.3.

Each simplex in figure 4.3 shows the case for a different value of topological randomness $q$. The general trend is that the lowest values of $\bar{\kappa}_c$ are found in the centre of the simplex for all values of $q$, with the highest values around the edge of the simplex. The values of $\bar{\kappa}_c$ decrease with increasing $q$, meaning that more random networks are more conducive to synchronisation. We have not yet investigated the influence of the topology of the micro-grid models; we have only discussed the power variability of their nodes. The choice of micro-grid network topology will determine onto which simplex from figure 4.3 the trajectory will be mapped. However, micro-grids are still an emerging technology lacking clear unifying topological features. We shall therefore pick two extreme cases for the underlying topology of the micro-grid models; those of regular lattices with $q = 0$ and of Poisson networks with $q = 1$. We shall also pick two seasonal extremes: winter and summer. Two trajectories, representing a week in winter and a week in summer respectively, will therefore be projected onto both the $q = 0$ and $q = 1$ simplexes from figure 4.3. This will reveal the variation in $\bar{\kappa}_c$ over the course of an average week in summer and an average week in winter, for two different underlying topologies.

The results are shown in figure 6.5. All cases are for micro-grids of size $n = 50$ with 100% PV uptake. Ensemble sizes are $\mathcal{N} = 50$. Panel 6.5(a) shows the ensemble mean trajectory for a week in winter overlaid onto the $\bar{\kappa}_c$ simplex for regular lattices, taken from the data for figure 4.3(a). The corresponding $\bar{\kappa}_c$ time-series is displayed beneath and shows that $\bar{\kappa}_c = 0.5$ for most of the week. This value of $\bar{\kappa}_c$ corresponds to times where the trajctory is adhered to either of the two lateral edges of the simplex. We know from section 6.2.4 that these points occur during night time when grid activity is dominated by consumption, and around midday when activity is dominantly productive. Each day there are two drops in $\bar{\kappa}_c$. These dips correspond to times when the net activity swings from production to generation and back again, causing the trajectory to cross through the centre of the simplex. This occurs in the morning when demand drops and PV increases, and in the afternoon when the opposite occurs. The same behaviour is visible in figure

6.5(b), which shows the same trajectory but where the underlying topology is Poissonian. The values of $\bar{\kappa}_c$ are lower in this case than for the lattice, which is to be expected as Poisson networks are more conducive to synchronisation. The bottom two panels, (c) and (d), show the equivalent results but for an average week in summer. Again, the same behaviour as in winter can be observed, but the daily dips are more abrupt. This is due to higher volumes of PV power output in the summer, causing faster and more pronounced swings across the simplex.

Taken together, the results of figure 6.5 show that the critical coupling capacity is in general briefly minimised during the day for a micro-grid, when production and supply switch back and forth in prominence. An effective way to minimise the critical coupling capacity, and thus ensure grid synchronisation, may be to try and force the system to spend longer periods occupying the centre of the simplex. This might be achievable through some combination of smart scheduling, pricing, and battery technologies.

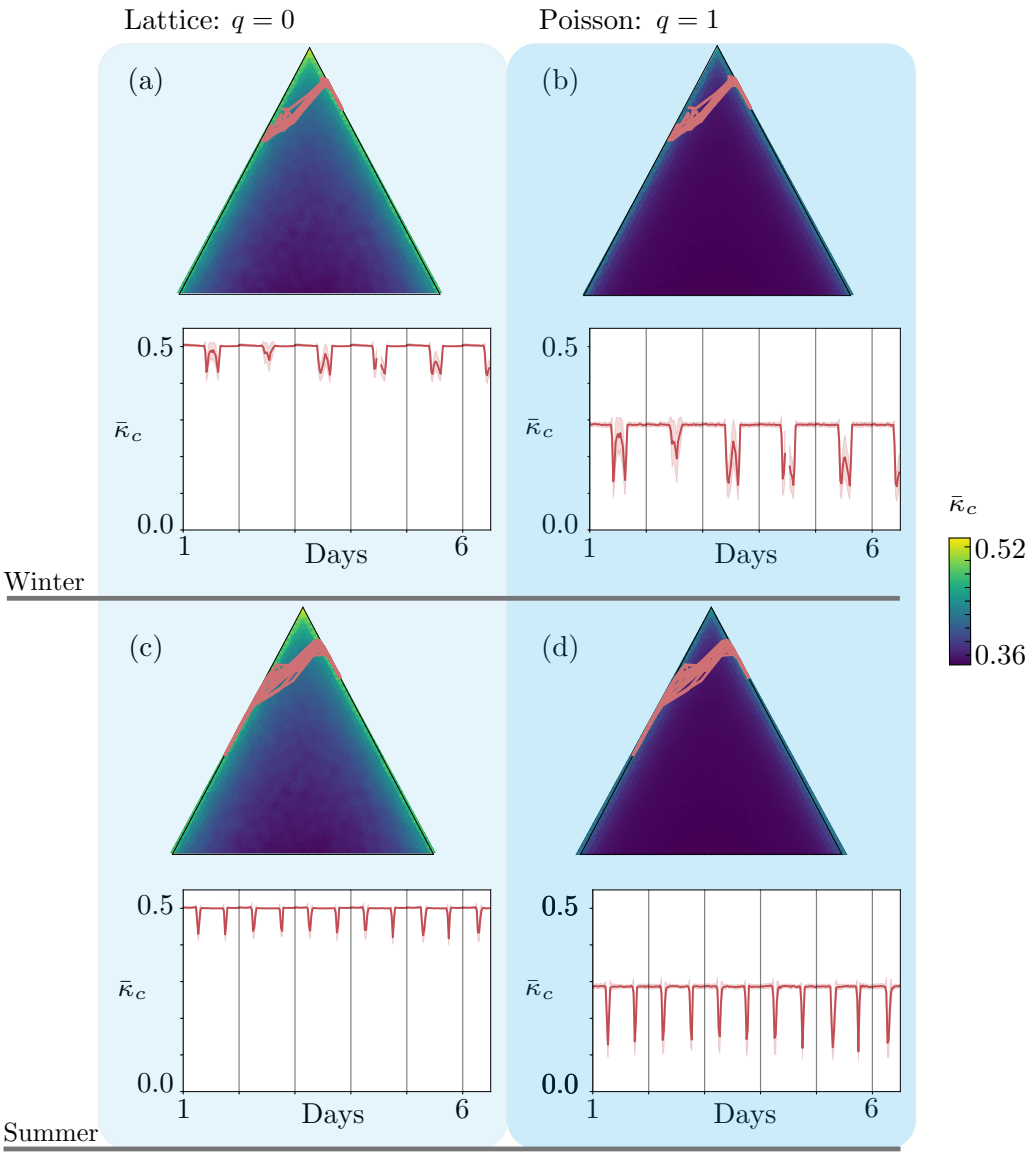



Figure 6.5: The mean variation in the critical coupling capacity $\kappa_c$ of a micro-grid over the course of a week, for four different scenarios. The colours in the simplex plots encode the mean critical coupling capacity $\bar{\kappa}_c$ as computed in chapter 4. Simplex labels have been omitted for visual clarity. The red lines in each simplex show the mean weekly trajectory of 50 micro-grid realisations each with 50 nodes. The time-series plots show the progression of $\bar{\kappa}_c$ as the trajectory progresses. Panel (a) shows the case of a week in winter where the underlying network structure is a regular lattice. Panel (b) shows the same trajectory but where the underlying network is Poisson. Panel (c) shows a week in summer for a regular lattice micro-grid, and (d) shows the same week on a Poisson network.

## 6.4 Variability in resilience

Chapters 3 and 5 investigated the resilience of electrical networks to cascading failures. Their resilience was found to vary as a function of the node composition of the networks, as well as with topological randomness $q$. This section investigates how the dynamic, mutable node composition of micro-grids impacts their resilience to cascading failures.

### 6.4.1 Methodology

As in the previous section, an ensemble of $\mathcal{N}$ micro-grids will be instantiated, each having $n$ houses. The trajectory of each grid in the ensemble will then be tracked though the $(\eta_+, \eta_-, \eta_p)$ simplex over the course of a week. To assess resilience, a cascading failure will then be triggered randomly at some point in the week. The time at which each cascade occurs will be chosen with a probability proportional to the maximum power in the network, in order to simulate the phenomenon that blackouts and failures occur during periods of high demand (Dobson et al., 2007). Specifically, let the maximum power in the network at time $t$ during a trajectory be denoted

$$P^{\max}(t) = \max |P(t)|. \tag{6.4}$$

$P \in \mathbb{R}^n$ is the power vector expressing the activity at time $t$ of each node in the micro-grid. The probability of a cascade being triggered at $t$ is then chosen to be

$$p(t) = \frac{P^{\max}(t)}{\sum_{t'} P^{\max}(t')}, \tag{6.5}$$

where the denominator is the sum of all of the maximum powers over the week-long trajectory.

From chapter 5, it is known that the resilience of a power network to a cascade will depend upon its node composition. In that chapter, and in chapter 3, resilience was measured using a metric $\rho$. This is defined as the ratio of the critical edge capacity $\alpha_c$ of a network required for at least half of

the network structure to survive a cascade, to the maximum flow volume $\alpha_*$ in the network at the start of the cascade:

$$\rho = \frac{\alpha_c}{\alpha_*}. \tag{6.6}$$

The lower the value of $\rho$, the less capacity is required to survive a cascade and so the more inherently resilient the network composition.

The resilience of a micro-grid to a cascade will depend upon its location in the simplex at the time when the cascade is triggered. To gauge this, the coordinate $(\eta_+, \eta_-, \eta_p)$ of the cascade point will be multiplied by the number of nodes $n$ to get effective source and sink numbers. Recall from chapter 5 that the mean resilience $\bar{\rho}$ was computed in the simplex for networks of sizes $n = 50$ and $n = 100$, for various values of topological randomness $q$. Recall also that for a given point in the simplex, the value of $\rho$ is log-normally distributed. The parameters for the underlying log-normal fits for the simplex data in chapter 5 provide a computationally efficient way of simulating a cascade; all that is needed is the appropriate log-normal parameters for the given cascade point in the simplex. From this, the log-normal distribution can be sampled, to provide a value of resilience for that point in the simplex for a random realisation of a Watts-Strogatz network with given $q$ and $n$. The critical capacity required to survive a cascade can then be computed as $\alpha_c = \rho\alpha_*$. The lower the value of $\alpha_c$, the more resilient the network. $\alpha_*$ is obtained from the data using a steady-state calculation. Specifically, the data-fed power vector $P$ is used to compute the steady state flow pattern using the linearised DC flow approximation from Chapter 3. The maximum flow volume in this steady state is $\alpha_*$. Consequently, these resilience experiments use an efficient hybrid steady-state dynamical method; the steady state is used for computing $\alpha_*$, while the sampling method allows the swing-equations dynamics to be taken into account for the cascades which occur on a fast time-scale.

The above procedure for computing $\alpha_c$ will be repeated for large ensembles of micro-grid realisations and provides a computationally efficient means of characterising resilience to cascades, without having to simulate a cascade directly. As with the previous chapter, micro-grids will be chosen to have

size $n = 50$ and various different test cases of $q$, season and uptake will be considered.

### 6.4.2 Results

All ensemble sizes in this section are $\mathcal{N} = 500$. Each micro-grid in each ensemble has 50 nodes and is assumed to have a Watts–Strogatz underlying topology with mean degree $k = 4$, and parameterised by topological randomness $q \in [0, 1]$. Figure 6.6 shows the results for an ensemble in the winter whose underlying topologies are lattices ($q = 0$). Panel 6.6(a) shows the case where each micro-grid has only 50% PV uptake. The simplex in (a) shows the variation of $\bar{\rho}$ with node composition for lattices with $n = 50$, as computed in chapter 5. The dots in the simplex show the locations of each cascading failure. Most occur at the upper right edge of the simplex, corresponding to times when there is little PV production and the micro-grids are dominated by heavy demand in the evenings. We know from section 6.2.4 that the 50% uptake case in the winter spends most of the day in this region of the simplex. The histogram in panel (c) shows the corresponding values of the the critical edge capacity $\alpha_c$ required to survive each of the cascades. The mean value is $\bar{\alpha}_c = 4.6$kW. This means that on average, a micro-grid with 50% PV uptake must have connections rated to carry at least 4.6kW in order for most micro-grids to survive a potential cascading failure. Panel (b) shows the case for 100% uptake, and reveals that roughly half of the cascades happen in the upper right when the gird is dominated by demand and the other half on the left hand side when the grid is dominated by PV. A few cascades also occur towards the centre of the simplex. Panel (d) shows the corresponding critical capacities, with a mean value of $\bar{\alpha}_c = 4.5$kW. This is roughly 0.1kW less than in the 50% uptake case. This means that in the winter, and if the grid topology is regular, then increasing the proliferation of PV on the grid will bring only a small increase in resilience.

Figure 6.7 shows the case of lattice micro-grids in the summer. Panel (a) shows that the 50% PV uptake case leads to cascades occurring on both the right and left side of the simplex. Compared to the 50% case in winter, shown

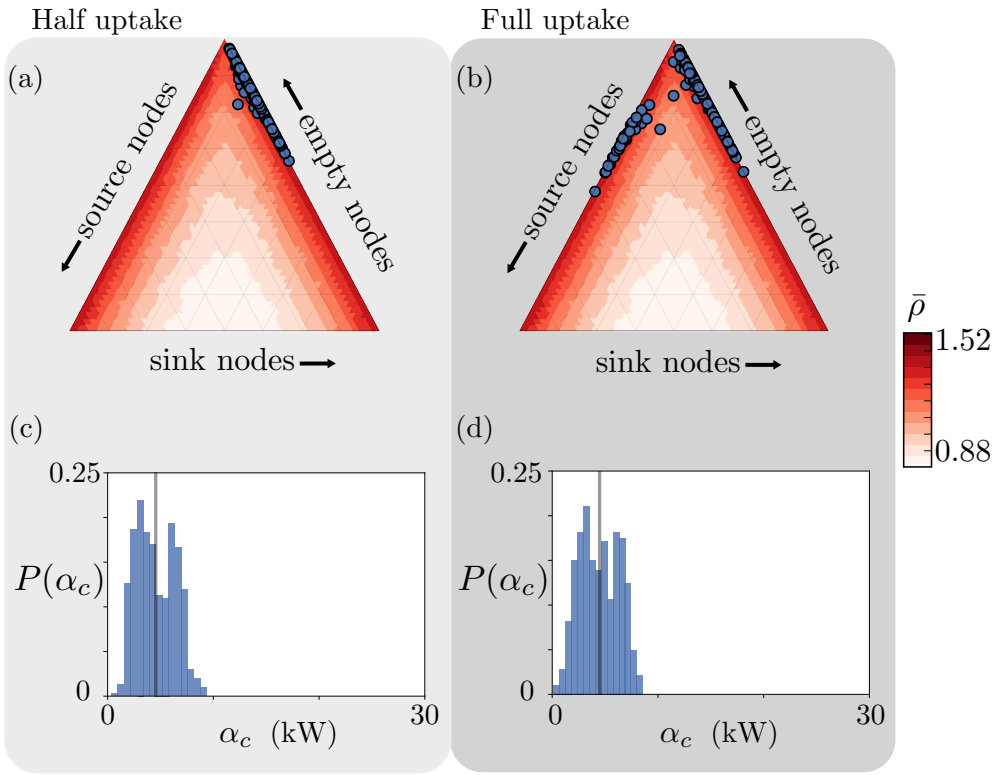



Figure 6.6: Random cascading failures occurring during a week in winter for an ensemble of 500 micro-grid realisations, where the underlying network topologies are all lattices ($q = 0$). The procedure used to generate these results is outlined in section 6.4.1. Each micro-grid has 50 nodes. The values of resilience $\bar{\rho}$ displayed within each simplex were computed in chapter 5. The dots in each simplex indicate the locations where cascading failures occurred. The histograms beneath each simplex show the corresponding values of edge capacity $\alpha_c$ required to survive each cascade. The mean $\bar{\alpha}_c$ is indicated by the vertical line in each histogram. The higher the value of $\bar{\alpha}_c$, the less resilient the micro-grid. For the half PV uptake case in the left hand column, $\bar{\alpha}_c = 4.6$kW. For the full uptake case in the right column, $\bar{\alpha}_c = 4.5$kW.

in 6.6(a), the summer cascades have a higher tendency to occur in the left side of the simplex. This is indicative of the higher PV production in the summer, meaning that even in the 50% PV case, there are times of the day when the grid is dominated by production. Figure 6.7(c) shows the corresponding values of critical capacity, with a mean of $\bar{\alpha}_c = 4.1$. This is less than in the winter case; the higher volumes of PV cause more cascades to occur in the more resilient upper central region of the simplex. The case for 100% PV uptake in the summer for lattice micro-grids is shown in panel (b). Due to the

very high volumes of PV production in such grids in the summer, the daily trajectories are able to travel far down the left hand side of the simplex, into regions dominated by high source term concentrations. Consequently, many cascades happened in this region. The values of $\bar{\rho}$ increase down the sides of the simplex where these cascades are occurring. The values of $\alpha_c$ are therefore higher, as shown in panel (d). The mean is $\bar{\alpha}_c = 10.3$kW, higher than any other case considered thus far. Having a high PV uptake can therefore cause a lack of resilience in summer months, due to the high volumes of power they output at midday which must all be shunted to the PPC since there will be insufficient demand within the network to consume it. It is also notable

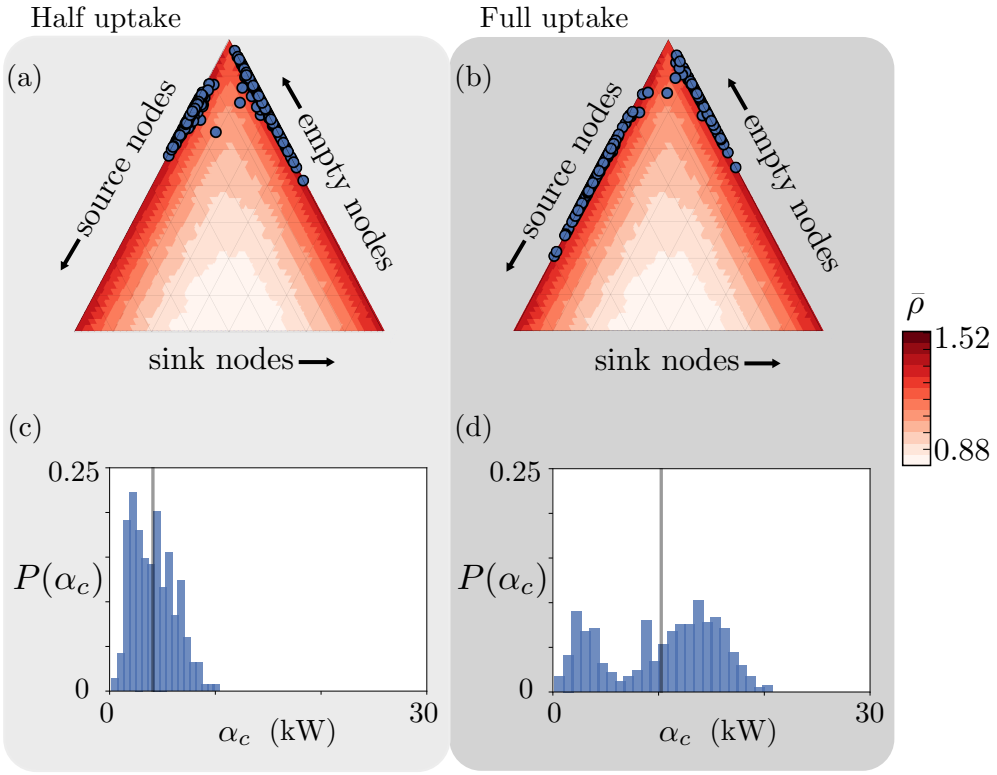



Figure 6.7: Random cascading failures occurring during a week in summer for an ensemble of 500 micro-grid realisations, where the underlying network topologies are all lattices ($q = 0$). Each micro-grid has 50 nodes. The dots in each simplex indicate the locations where cascading failures occurred. The histograms beneath each simplex show the corresponding values of edge capacity $\alpha_c$ required to survive each cascade. For the half PV uptake case in the left hand column, $\bar{\alpha}_c = 4.1$kW. For the full uptake case in the right column, $\bar{\alpha}_c = 10.3$kW.

that the distribution $\alpha_c$ for the 100% uptake case in the summer is bimodal. The smaller mode corresponds to cascades occurring in the upper right of the simplex. These are cases which occur in the evening when the grid is dominated by heterogeneous demand, and yield less damaging cascades. The larger mode corresponds to the cascades occurring during high PV output.

Figures 6.8 and 6.9 show the equivalent results when the underlying network topology is Poisson ($q = 1$). They show the same qualitative behaviour as in the lattice cases of the previous two figures. Namely, that the resiliences of both the 50% and 100% uptake cases are approximately the same during the winter, and that during the summer the 100% uptake case shows a relatively high vulnerability to cascades due to high volumes of PV as shown in 6.9(d). The values of $\alpha_c$ are slightly higher across the board for Poisson networks, which is to be expected since chapter 5 showed that Poissonian networks were less resilient than lattices.

The values of the critical coupling capacity found in this section are mostly within the approximate range 5 to 15kW. In low voltage distribution grids, such as micro-grids, the network connections are typically armoured low voltage cables. These cables tend to be rated for maximum power flows in the range 4 to 15kW (British Standard, 2018), depending on cable thickness. These two ranges overlap, meaning that real-world distribution scale grids are operating at close to the critical coupling capacity $\alpha_c$. In summary, this section has shown that increasing the volume of PV can have an adverse effect on resilience during summer months, due to the high volumes of power that must be handled by the network.

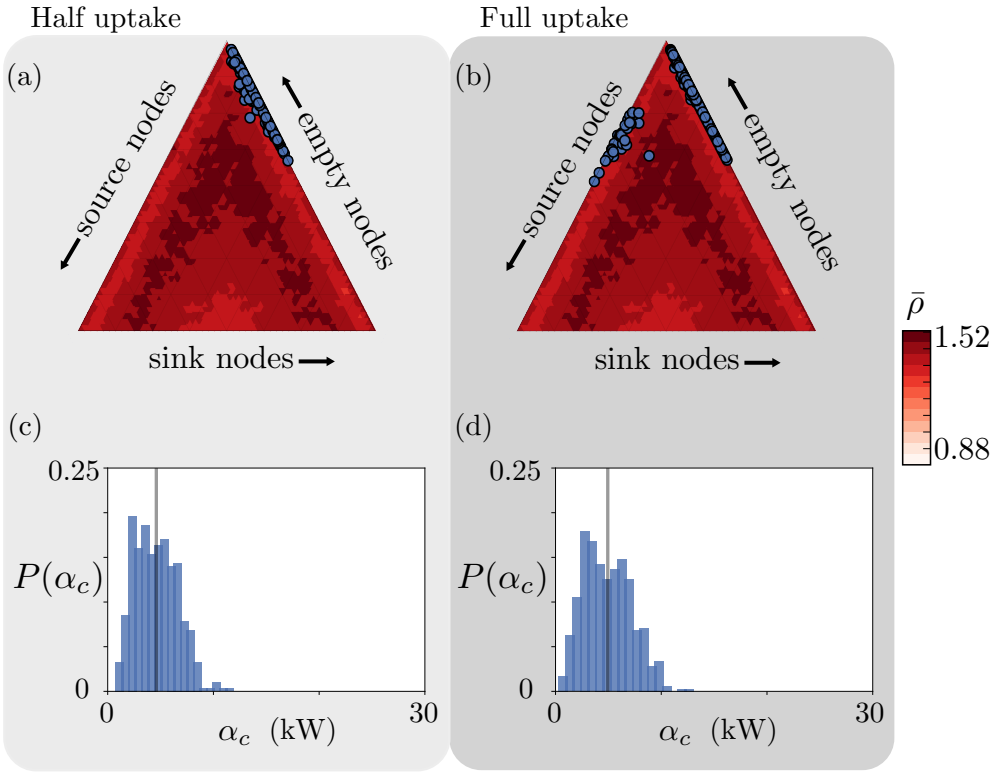



Figure 6.8: Random cascading failures occurring during a week in winter for an ensemble of 500 micro-grid realisations, where the underlying network topologies are all Poisson $(q = 1)$. Each micro-grid has 50 nodes. The values of resilience $\bar{\rho}$ displayed within each simplex were computed in chapter 5. The dots in each simplex indicate the locations where cascading failures occurred. The histograms beneath each simplex show the corresponding values of edge capacity $\alpha_c$ required to survive each cascade. For the half PV uptake case in the left hand column, $\bar{\alpha}_c = 4.6$kW. For the full uptake case in the right column, $\bar{\alpha}_c = 4.9$kW.

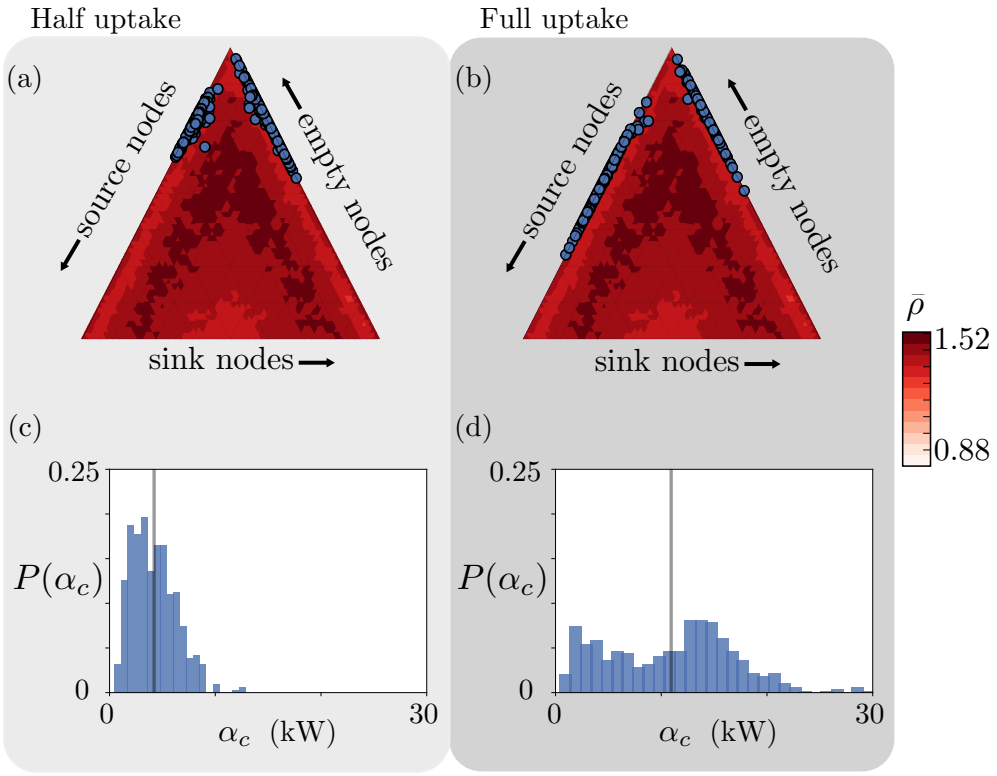



Figure 6.9: Random cascading failures occurring during a week in summer for an ensemble of 500 micro-grid realisations, where the underlying network topologies are all Poisson ($q = 1$). Each micro-grid has 50 nodes. The dots in each simplex indicate the locations where cascading failures occurred. The histograms beneath each simplex show the corresponding values of edge capacity $\alpha_c$ required to survive each cascade. For the half PV uptake case in the left hand column, $\bar{\alpha}_c = 4.2$kW. For the full uptake case in the right column, $\bar{\alpha}_c = 10.9$kW.

## 6.5 Effect of batteries

Many micro-grid designs incorporate batteries. These are used to smooth the disparity between supply and demand and thus increase the self-sufficiency of the grid. PV power produced during the day can be stored for periods of high demand in the evening. This chapter uses a simple model of a household battery to investigate time impact of storage upon the trajectories of micro-grids in the node configuration simplex, and their resilience.

It will be assumed that each house in the grid that has PV also has a lithium ion battery. As a reference, the battery model will be based on the Tesla Powerwall 2 (Tesla, 2018). This is a type of lithium ion battery commonly being installed into sustainable housing projects. It can charge and discharge at a maximum rate of 5kW and has a maximum capacity of 14kWh. The behaviour of the batteries will be modelled as follows. At each point in time, the power demand of a given house and its PV power output will be measured. If production exceeds demand, the excess power will be used to charge the

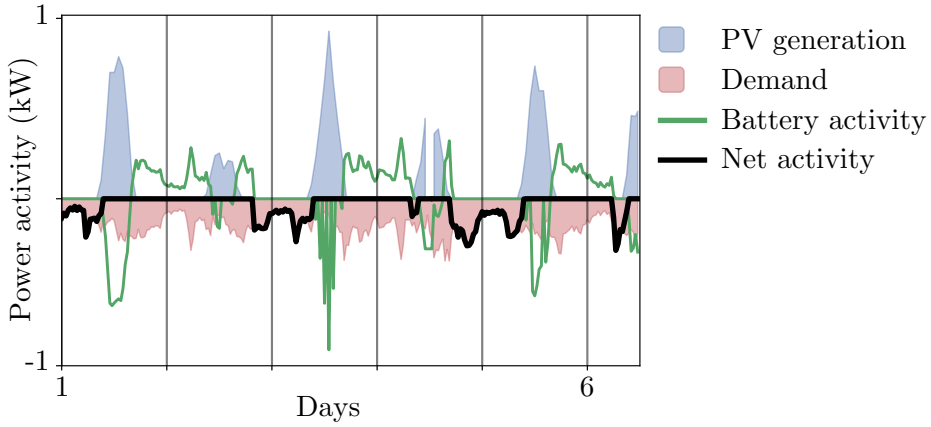



Figure 6.10: The power activity of a house equipped with both PV and a battery, over the course of a week in January. Negative values indicate consumption. Each vertical line represents midnight each day. The blue and red shaded regions indicate PV generation and household power consumption respectively. The green line indicates battery activity. When the line is positive, the battery is discharging power to be consumed by the household. When it is negative, the battery is consuming and storing power from the PV unit. The black line indicates the net power demand of the household. When it is zero, the house's demand is matched by its own supply.

battery at up to the maximum charging rate. Any further excess will be supplied back to the grid as usual. Similarly, if demand exceeds supply, the battery will discharge at up to its maximum rate. If this is insufficient, any remaining power demand will be drawn from the grid. The net power activity of such a house is thus depends on three terms: the battery activity, the PV production, and the power demand. Figure 6.10 shows each of these values over the course of a week for a such a house. The black line represents the net power of the house. The figure shows that the battery is able to store sufficient power during times of PV production to cover some periods of high demand occurring later in each day. In such situations, the net activity shown by the black line is zero. This means the house is self-sufficient during these periods. If the net activity is positive, then the house is effectively a generator and will contribute to the generation density term $\eta_+$ when computing the grid's node composition. Similarly, a net value less than zero will contribute to $\eta_-$. It is to be expected that batteries will increase the passivity $\eta_p$ of the grid since more houses will net zero and thus draw little from the grid.

Figure 6.11 shows the mean weekly trajectories through the node configuration simplex for battery equipped micro-grids. The figure shows the case for 50% and 100% PV uptake in both summer and winter. During winter, shown in (a) and (b), both uptake cases live entirely on the upper right edge of the simplex and do not travel into the simplex interior at all. This means that the grids are dominated entirely by small, heterogeneous sink nodes. This is due to the relatively low PV outputs of the winter months being small enough to be completely stored within household batteries for later consumption. The result is no net generators, since all houses consume their own PV power entirely. During the summer months, shown in panels (c) and (d), there is much more PV power on grid. Battery capacity is insufficient to store all of this PV power, leading to houses becoming net generators to the grid. As a result, the trajectories take excursions to the other side of the simplex. In comparison with the equivalent cases in 6.2.4 that had no batteries, the trajectory of the 100% uptake summer case is still more passive and tends to hug the exterior of the simplex.

The impact of batteries on the critical coupling capacity $\kappa_c$ is shown in

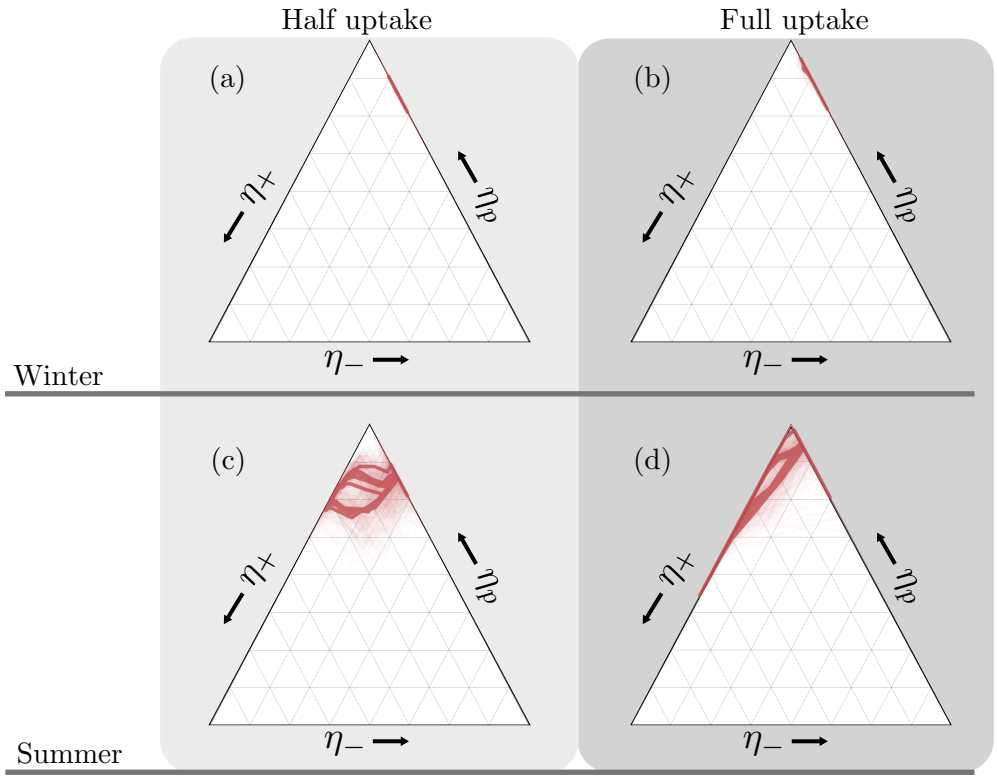



Figure 6.11: The thick red lines indicate the mean weekly trajectories for ensembles of 50 micro-grids with $n = 50$ nodes. The left hand column represents networks where only half of the houses have PV panels, and the right where all houses have PV. Each row represents a different season. Each house that has PV panels is also equipped with a battery.

figure 6.12. Refer back to section 6.3 for details of the computation of $\kappa_c$. Recall that the mean critical coupling capacity $\bar{\kappa}_c$ was found to exhibit daily decreases that correspond to points where the grid transitions from being net generative to net consumptive. Panels 6.12(a) and (b) show that these daily decreases are no longer present during the winter in grids with batteries. This is because there is now insufficient PV power to produce any net generators, and therefore there are no excursions across the simplex. During the summer, as shown in panels (c) and (d), the daily decreases are still visible. However, they are smaller and more abrupt than in the non-battery case. This is because the excursions across the simplex in the summer are quicker, due to battery activity restricting the period of time each day that PV is generating enough to cause net generation.

The impact of batteries on the resilience of micro-grids to cascades is shown in figure 6.13. This figure focuses on the 100% PV penetration during the summer, since this scenario was found in 6.4.2 to cause the greatest vulnerability to cascades. We now wish to see if batteries can improve the situation. Figure 6.13 shows that the inclusion of batteries has caused the cascades to occur during configurations that are closer to the edge of the simplex. The resulting distributions of the critical edge capacity $\alpha_c$ are bimodally distributed, as in the non-battery case, corresponding to the cascades occurring during either times of net production or net consumption. The mean values of critical capacity $\bar{\alpha}_c$ in the lattice and Poisson case are $\bar{\alpha}_c = 9.5$kW and $\bar{\alpha}_c = 10.3$kW respectively. These values are approximately 1kW less than the equivalent non-battery cases analysed in 6.4.2, meaning that the batteries have caused an approximately 10% decrease in the mean capacity required to survive a cascade in the summer.

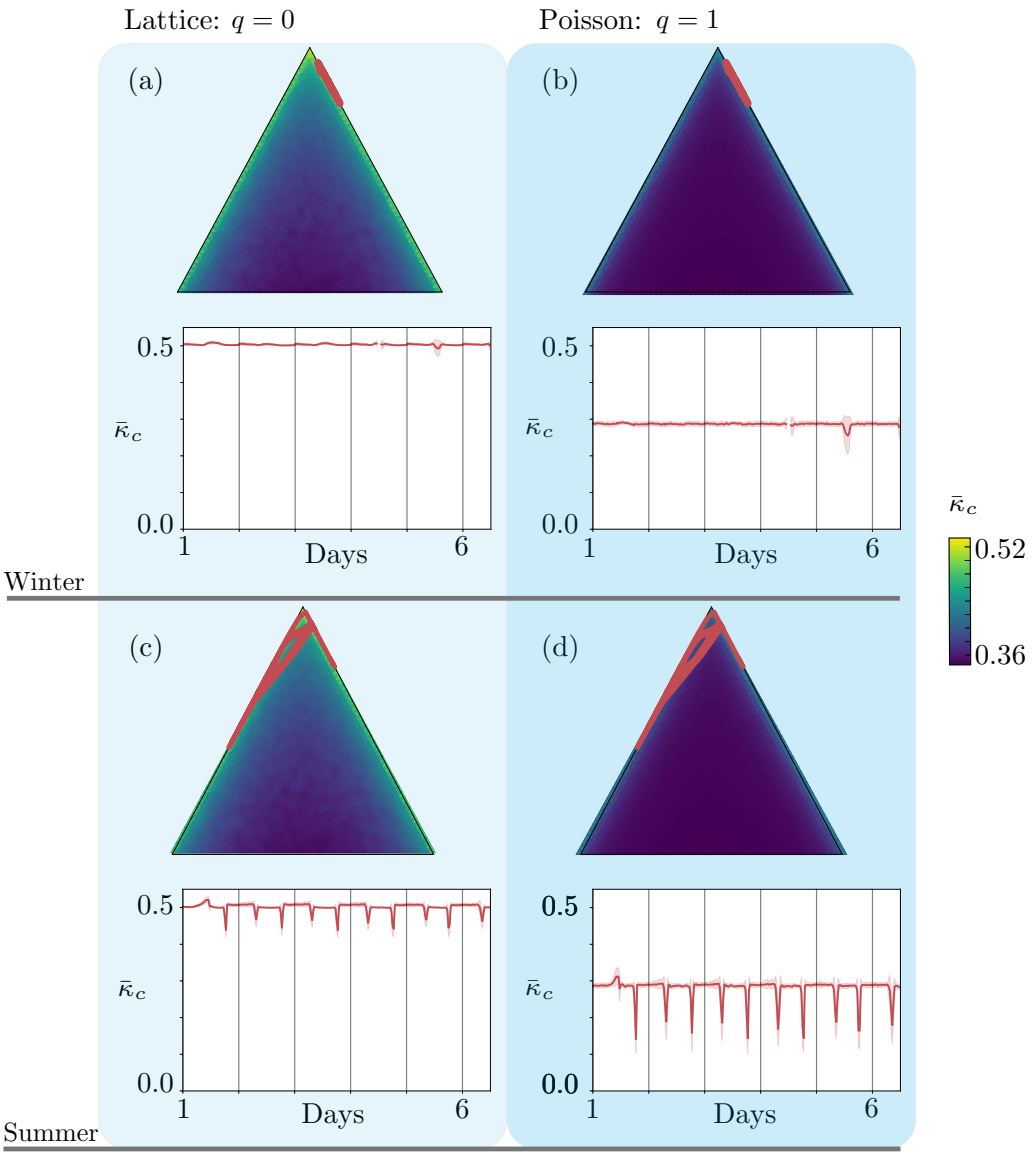



Figure 6.12: The mean variation in the critical coupling capacity $\kappa_c$ of a micro-grid over the course of a week, for four different scenarios. Each household is equipped with a battery. The colours in the simplex plots encode the mean critical coupling capacity $\bar{\kappa}_c$ as computed in chapter 4. The red lines in each simplex show the mean weekly trajectory of 50 micro-grid realisations each with 50 nodes, each node having an associated battery. The time-series plots show the progression of $\bar{\kappa}_c$ as the trajectory progresses. Panel (a) shows the case of a week in winter where the underlying network structure is a regular lattice. Panel (b) shows the same trajectory but where the underlying network is Poisson. Panel (c) shows a week in summer for a regular lattice micro-grid, and (d) shows the same week on a Poisson network.

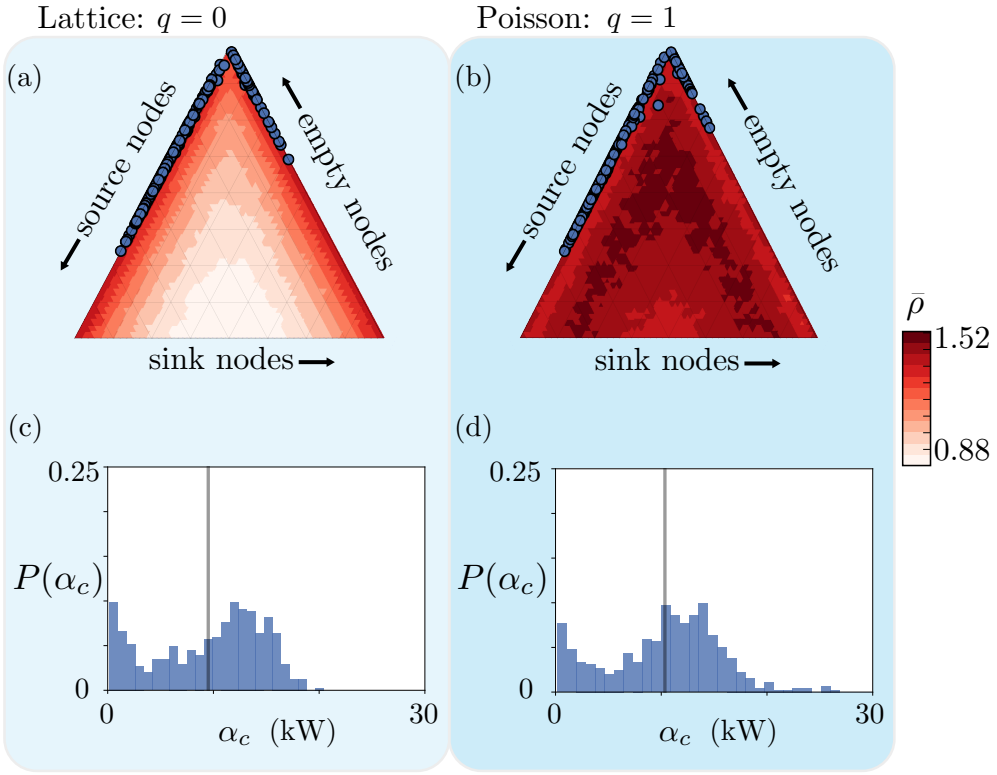



Figure 6.13: Random cascading failures occurring during a week in summer for an ensemble of 500 micro-grid realisations during the summer. Each micro-grid has complete battery and PV uptake. The left hand column shows the case where the underlying network structure is a regular lattice. The right hand column shows the case where is Poisson. Each micro-grid has 50 nodes. The dots in each simplex indicate the locations where cascading failures occurred. The histograms beneath each simplex show the corresponding values of edge capacity $\alpha_c$ required to survive each cascade. For the lattice case in the left hand column, $\bar{\alpha}_c = 9.5$kW. For the Poisson case in the right column, $\bar{\alpha}_c = 10.3$kW.

## 6.6 Conclusion

This chapter introduced a model of electrical micro-grids. These are low voltage, household distribution scale power grids comprising some tens of houses. They tend to incorporate renewable technologies such as batteries and photo-voltaic (PV) panels. Real household level power consumption and PV generation data were incorporated into the micro-grid models to produce a realistic framework within which to analyse the daily and seasonal variability of renewable generation.

The effective node composition of a micro-grid is mutable and changes throughout a day. At some times, all of its nodes may function as consumers; at other times the nodes function as generators. The objective of the chapter was to investigate how this dynamic variability in node composition impacts key network features relevant to power grids; namely, critical coupling strength $\kappa_c$ and critical edge capacity $\alpha_c$. A method was devised to project the state of a micro-grid onto a node configuration simplex, such as those used in previous chapters. This then allowed the data from previous chapters to be sampled, revealing how $\kappa_c$ and $\alpha_c$ vary during the daily and weekly trajectory of a micro-grid through the node configuration simplex.

It was shown that the daily and weekly trajectories of micro-grids occupy the upper half of the node configuration simplex. This is to be expected, given their highly passive nature. The grids spend most of their time on either of the two lateral edges of the simplex, corresponding to whether they are dominated by generation or consumption. During the day, they take an excursion across the simplex and back again as the balance between supply and consumption flips.

The resilience of grids to cascading failures was investigated for various scenarios. Increasing the volume of PV in domestic settings is generally regarded as a means of building resilience and self-sufficiency. However the findings of this chapter highlight an unexpected drawback of PV panels; they can decrease resilience during the summer months. This results from the high volume of solar output that the grid must deal with in times where

demand is low. The introduction of batteries was shown to partly mitigate this seasonal vulnerability.

The critical coupling strength $\kappa_c$ required for grid synchronisation was found to decrease periodically each time the trajectory crossed the simplex. These decreases are advantageous since they boost the grids inherent ability to maintain synchrony. However, the introduction of household batteries was found to partially suppress these decreases.

This chapter has found that adding batteries and PV to grids does not have a straightforward impact. Although they certainly boost self-sufficiency, they can also negatively impact resilience and the grids ability to synchronise. This highlights the need to find more sophisticated control methods that may help to drive the trajectories of micro-grids further into the centre of the node configuration simplex, while still supporting self-sufficiency.

# Chapter 7

# Conclusion and further work

This thesis has investigated the resilience and efficiency of networks whose nodes have heterogeneous behaviour. The nodes may variably function as flow sources or sinks, and possess different abilities to adjust their flow in reaction to network conditions. In this final chapter, the contributions of this thesis to the understanding of resilience and efficiency will be summarised and areas of potential future work will be identified.

## 7.1 Network efficiency and redundancy

### 7.1.1 Summary

Chapter 2 investigated the efficiency of flow networks with heterogeneous nodes using the Price of Anarchy $\mathcal{P}$. Conventionally, $\mathcal{P}$ is used in game theory to study the impact of non-cooperation. Chapter 2 applied $\mathcal{P}$ to study flows such as electricity in networks by exploiting the equivalence between least-resistance flows and the Nash equilibria of traffic routing games. These Nash equilibria were shown to be solutions to Kirchoff's laws. $\mathcal{P}$ was then shown to provide a computational efficient measure of the inefficiency of a network flow.

$\mathcal{P}$ was computed as a function of the total flow volume in the network for each possible proportion of flow source, sink, and passive nodes. Each of the $n(n-1)/2$ possible node configurations was represented as a point on a triangular simplex. It was shown that for each possible configuration, $\mathcal{P}$ increases as a function of total flow volume and attains a single maximal peak. This peak indicates high inefficiency. The location and height of this peak for each node composition were then projected into the node configuration simplex. Remarkably the highest inefficiencies were found to occur when the numbers of source and sink nodes are equal; this is a configuration that may occur regularly in electrical grids with a high level of distributed generation. This phenomenon was observed across the full range of Watts–Strogatz networks, as well as in scale-free networks and real power grid topologies. The dependence of $\mathcal{P}$ on the configuration of node types was therefore found to be approximately invariant across network topologies. This was further verified by introducing various modes of noise and by varying network size to extract scaling behaviour. These scalings mean the results can be applied to networks of arbitrary size. The values of $\mathcal{P}$ were found to settle to a roughly constant value for each respective configuration as size $n$ increases beyond $n \approx 100$.

Chapter 2 also introduced a numerical algorithm for computing redundancy $\mathcal{R}$ in networks. $\mathcal{R}$ was defined as the percentage increase in efficiency available by capping the capacity of a single edge in the network. $\mathcal{P}$ was shown to be a proxy for $\mathcal{R}$. Crucially, the locations of peaks in $\mathcal{P}$ and $\mathcal{R}$ coincide. $\mathcal{P}$ therefore provides a computationally efficient means of assessing topological redundancy.

The results of chapter 2 provide a simple means of viewing inefficiency as a function of the proportion of flow source and sink nodes. This allows regions of high efficiency in the operating space to be located. In networks such as renewable grids, node types are mutable and can even be influenced by a central controller. Such a network could therefore be guided towards regions of high efficiency. However, since efficiency is tied to a lack of redundancy, and redundancy is allied to resilience, there is a trade-off to be made between efficiency and resilience. Indeed, subsequent chapters go on to show that the

regions typified by inefficiency are also regions that are resilient to cascading failure.

### 7.1.2 Further research

The results of chapter 2 present several open questions that might form the basis of future research. Firstly, chapter 2 allocated the locations of source and sink nodes at random in networks. It would be of interest to investigate how specific source-sink placement strategies on a given fixed network topology influence $\mathcal{P}$. It should be noted that Roughgarden (2006) found that designing networks to minimise $\mathcal{P}$ is NP-hard. Nonetheless, general principles of source-sink placement may still prove useful. For example, perhaps placing sources in highly central positions might in general decrease $\mathcal{P}$.

An obvious further question is how nonlinear network edge cost functions impact $\mathcal{P}$. It is possible that nonlinear functions lead to much higher values of $\mathcal{P}$. Additionally, the exact relationship between $\mathcal{P}$ and topological redundancy $\mathcal{R}$ should be investigated further. This thesis only showed that the two provide the same information about where the peaks of inefficiency with respect to flow volume occur. Providing an exact way of converting between the two measures would provide a much deeper insight.

A potentially fruitful line of research would be to investigate $\mathcal{P}$ on a multiplex network. This would have application to transportation systems. The two edge, two node Pigou network used repeatedly in chapter 2 to demonstrate $\mathcal{P}$ could in fact be regarded as a trivial two-layer multiplex; one layer having a variable cost and the other having a fixed cost. A general investigation into more complex two-layer systems may yield valuable insights into the efficiency of multi-mode transportation systems such as those in London, where travellers may use the bus layer of the underground rail layer.

## 7.2 Network resilience

### 7.2.1 Summary

While the analysis in chapter 2 focused on network efficiency, with application to both electrical and traffic flow networks, chapters 3 onwards focus specifically on the resilience of electrical networks. These chapters used a combination of numerical simulations and, in chapter 6, real world power consumption data to investigate how the resilience of electrical networks depends upon the proportions of source (generator) and sink (consumer) nodes.

Chapter 3 used a steady state model of power flow together with an iterative model of cascading network failures to assess resilience. For a given network, the fraction of nodes surviving a cascade was computed a function of the network's edge capacity. The critical edge capacity at which at least half of the network survives was used a measure of resilience. The lower the value of $\alpha_c$, the less excess capacity the network requires to survive and thus the higher its resilience. As in chapter 2, each of the network's nodes may be a source or sink, with each $n(n-1)/2$ possible configuration being represented as a unique point on the node configuration simplex. The value of $\alpha_c$ was then computed as a function of the networks node composition. This revealed that for regular lattices and small-world networks, an even mixture of source and sink nodes provides a high resilience to cascading failures. These compositions correspond to those that also confer high redundancy, as highlighted in chapter 2. For Poisson networks, the opposite was found to be true; higher distribution of sources and sinks yields lower resilience. The point of transition between these two scenarios was tracked as a function of topological randomness $q$ and network size $n$ for ensembles of random networks. This reveals whether, for a given network, a more equal mixture of source and sink terms will cause an increase or decrease in resilience. In addition, the distribution of the resilience metrics was identified to be log-normal across all Watts–Strogatz networks.

Chapter 4 introduced the swing equation, a dynamic model of power flow

capable of capturing the essential frequency dynamics of AC power grids. The critical coupling strength $\kappa_c$ required for the power grid to synchronise was computed as a function of node composition and found to be minimsed when the proportions of source and sink nodes are equal. This means that increased distribution of generation yields a power network with a greater inherent ability to support synchronous operation, an essential requirement for a power grid.

Chapter 5 then used the swing equation to develop a nonlinear dynamical model of a cascading failure. This differed from the cascade models of chapter 3, as this dynamical model was capable of capturing line failures caused by de-synchronisaton in addition to solely line overloads. De-synchronisations were shown to occur more frequently in networks with lower levels of distributed generation. Cascades were shown to evolve on a similar timescale to transient behaviours identified in chapter 4. The transients therefore play an important role in cascade dynamics, highlighting the need for a dynamical model such as the swing equation. Measuring the duration of cascades also revealed that the most damaging events occur the quickest. Catastrophic network destroying failures unfold over the course of just several seconds, meaning that strategies to avoid cascades must be preventative.

Computing the resilience of networks to these dynamic cascades as a function of node composition revealed that the greatest resilience for a regular network occurs when there is an equal mix of sources and sinks. The resilience then decreases as the randomness of the underlying network topology increases. These results are similar to those found using the steady state cascade model, but differ in some key respects. In particular, the inversion behaviour observed in chapter 3 is less pronounced and increasing the mean degree of the network is not as effective at boosting resilience. Finally, chapter 6 applied the resilience results of the preceding chapters to a data-driven model of electrical micro-grids. The chapter then analysed impact of increased solar power and network variability on resilience. It was found that while increasing the uptake of domestic photo-voltaic panels can lead to greater self-sufficiency, it can also lead to a decrease in resilience against power failures.

### 7.2.2 Further research

Much of the analysis of resilience in chapters 3 to 6 relied on synthetic random networks. This is for two reasons. First of all, the topological designs of renewable power networks such as micro-grids are still an area of ongoing research. The general structural design principles of these networks have yet to crystallise. Therefore a random, synthetic approach allows greatest generality. Secondly, for larger power grids, high quality topological data is hard to come by. Such data is typically of low resolution, with little detail about the fine network structure underpinning each node. These data sets are also typically badly maintained. The IEEE test networks are now 60 years old and yet still in common usage, such is the sparsity of good data. However, this data sparsity is beginning to change. Companies such as Facebook are now using a combination of ground and satellite imaging, together with machine learning to produce high detail maps of power infrastructure (Facebook, 2020). This sort of data could be invaluable in future research into electrical networks. A clear line of further work therefore stems from extending the analysis presented in this thesis into these complex and detailed data sets. For example, conducting the analysis of chapter 6 on such larger scale networks would likely result in less passive systems that trace greater trajectories through the node configuration simplex.

It was mentioned in section 1.5.2 that renewable sources of energy are connected onto power grids using control mechanisms such as phase-locked loops (Chung, 2000) or virtual synchronous machines (Chen et al., 2011). These technologies help the renewable source approximate the behaviour of conventional rotating generators, allowing the use of the swing equation in the analysis of networks that contain renewable sources. However, the control mechanisms themselves, in particular the virtual synchronous machine, are non-trivial are capable of producing various instabilities. An analysis of these control mechanisms and their interactions with the rest of the grid is outside of the scope of this thesis, however it forms ongoing research.

An important area of further work is to identify control strategies for scheduling the use of batteries in micro-grids. As shown in chapter 6, a naive usage

of batteries can in fact lead to an increase in the critical coupling strength required for synchronisation. This is due to the increased passivity that the batteries introduce. Devising a control scheme to utilise batteries in such a way as to boost self-sufficiency whilst also promoting dynamical stability would be valuable future work.

# Appendix I

The optimisation problems in (3.3) and (3.5) are both convex, quadratic objective functions of the form $\mathcal{C} : \mathbb{R}^m \to \mathbb{R}$, subject to a set of linear constraints $Ef = b$. The method of subgradient-projection (Boyd et al., 2004) is used to solve both of these systems. This is an iterative scheme where the approximation $f \in \mathbb{R}^m$ at the $k^{\text{th}}$ step is

$$f^{(k)} = \mathbb{P}\left(f^{(k-1)} - \gamma^{(k-1)} \nabla \mathcal{C}(f^{(k-1)})\right), \tag{A1}$$

where $\gamma^{(k-1)}$ is the step-size at iteration $k-1$ and $\mathbb{P} : \mathbb{R}^m \to \mathbb{R}^m$ is the Euclidean projection of a vector onto the feasible set $\{f | Ef = b\}$ given by

$$\mathbb{P}(y) = (I - E^T(EE^T)^{-1}E)\, y + E^T(EE^T)^{-1}b. \tag{A2}$$

The step-size is chosen to be non-summable but diminishing:

$$\gamma^{(k)} \geq 0, \qquad \lim_{k\to\infty} \gamma^{(k)} = 0, \qquad \sum_{k=1}^{\infty} \gamma^{(k)} = \infty, \tag{A3}$$

for which convergence of the subgradient-projection method is guaranteed (Boyd et al., 2004). The particular choice used here is $\gamma^{(k)} = 0.1/\sqrt{k}$. The projector in the subgradient method requires the inversion of the Laplacian $EE^T$, but this may be done ahead of time. Then at each iteration only two sparse matrix-vector multiplications are required. Thus this method is very fast, despite the high number of iterations typically involved. The inversion of the Laplacian is problematic since, for a simple connected graph with $n$ nodes, the incidence matrix will have rank $n-1$. This means $E$ and $EE^T$ are rank deficient and singular. To overcome this problem, two additional

'virtual' edges are added which are connected at one end to any arbitrary node, but are unconnected at the other end. This breaks the simplicity of the graph and adds two extra columns to $E$. An additional row must then be added to $E$ to demand that the extra edges always possess zero flow. In this way the rank of $E$ is made equal to its dimension without affecting the flow on the original network, and the resulting Laplacian $EE^T$ is well conditioned. The convergence of this method is shown in Figure A1.

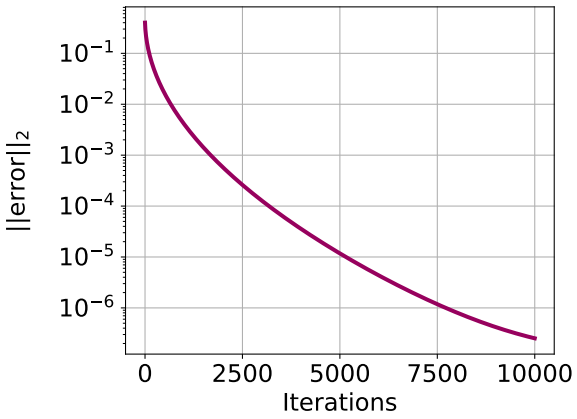



Figure A1: Convergence of the subgradient-projection method for finding the Nash flow on a square lattice network with 120 nodes. Note that despite the high number of iterations, the method is still fast since each iteration involves just a sparse matrix-vector multiplication.

# References

N. Ahmed, R.A. Michelin, W. Xue, S. Ruj, R. Malaney, S.S. Kanhere, A. Seneviratne, W. Hu, H. Janicke, and S.K. Jha. A survey of covid-19 contact tracing apps. *IEEE Access*, 8:134577–134601, 2020. doi: 10.1109/ACCESS.2020.3010226.

S. Albers, S. Eilts, E. Even-Dar, Y. Mansour, and L. Roditty. On nash equilibria for a network creation game. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithm*, pages 89–98, Philadelphia, PA, USA, 2006. SIAM. doi: 10.1145/2560767.

R. Albert and A. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, 2002. doi: 10.1103/RevModPhys.74.47.

R. Albert, H. Jeong, and A. Barabási. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000. doi: 10.1038/35019019.

A. Aleta, D. Martín-Corral, A. Pastore y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N.E. Dean, M.E. Halloran, I.M. Longini Jr, S. Merler, A. Pentland, A. Vespignani, E. Moro, and Y. Moreno. Modelling the impact of testing, contact tracing and household quarantine on second waves of covid-19. *Nat. Hum. Behav.*, 4(9):964–971, 2020. doi: 10.1038/s41562-020-0931-9.

C. Alexander and M. Sadiku. *Fundamentals of Electric Circuits*. McGraw-Hill Education, 2012. ISBN 0262182432.

C.N. Angstmann, I.C. Donnelly, and B.I. Henry. Pattern formation on networks with reactions: A continuous-time random-walk approach. *Phys. Rev. E*, 87:032804, 2013. doi: 10.1103/PhysRevE.87.032804.

APG. https://www.apg.at/en/Stromnetz, 2017. Accessed: 2019-05-17.

A. Arenas, W. Cota, J. Gomez-Gardenes, S. Gómez, C. Granell, J.T. Matamalas, D. Soriano-Panos, and B. Steinegger. A mathematical model for the spatiotemporal epidemic spreading of covid19. *medRxiv*, 2020. doi: 10.1101/2020.03.21.20040022.

M. Balcan, A. Blum, and Y. Mansour. Circumventing the price of anarchy: Leading dynamics to good behavior. *SIAM J. Comput.*, 42(1):230–264, 2013. doi: 10.1137/110821317.

Y. Bar-Yam. *Dynamics of complex systems*. CRC Press, 2019. doi: 10.1201/9780429034961.

A. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999. doi: 10.1126/science.286.5439.509.

A. Barabási et al. *Network science*. Cambridge university press, 2016.

BBC. Coronavirus: UK confirms plan for its own contact tracing app. https://www.bbc.co.uk/news/technology-52263244, 2020. Accessed: 2020-11-13.

A.R. Bergen and D.J. Hill. A structure preserving model for power system stability analysis. *IEEE Trans. Power App. Syst.*, (1):25–35, 1981. doi: 10.1109/tpas.1981.316883.

G. Bianconi and A. Barabási. Competition and multiscaling in evolving networks. *EPL*, 54:436, 2001. doi: 10.1209/epl/i2001-00260-6.

P. Bonacich. Some unique properties of eigenvector centrality. *Social Networks*, 29(4):555 – 564, 2007. doi: 10.1016/j.socnet.2007.04.002.

J. Borge-Holthoefer and Y. Moreno. Absence of influential spreaders in rumor dynamics. *Phys. Rev. E*, 85:026116, 2012. doi: 10.1103/PhysRevE.85.026116.

S. Bornholdt and T. Rohlf. Topological evolution of dynamical networks:

Global criticality from local dynamics. *Phys. Rev. Lett.*, 84(26):6114, 2000. doi: 10.1103/physrevlett.84.6114.

S. Boyd, S.P. Boyd, and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

D. Braess, A. Nagurney, and T. Wakolbinger. On a paradox of traffic planning. *Transp. Sci.*, 39(4):446–450, 2005. doi: 10.1287/trsc.1050.0127.

U. Brandes. A faster algorithm for betweenness centrality. *J. Math. Sociol.*, 25(2):163–177, 2001. doi: 10.1080/0022250X.2001.9990249.

British Standard. *Requirements for Electrical Installations, IET Wiring Regulations, Eighteenth Edition, BS 7671:2018*. BSI, 2018.

C.D. Brummitt, R.M. D'Souza, and E.A. Leicht. Suppressing cascades of load in interdependent networks. *Proc. Nat. Acad. Sci. USA*, 109(12): E680–E689, 2012. doi: 10.1073/pnas.1110586109.

S. Buldyrev, R. Parshani, G. Paul, H. Stanley, and S. Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464:1025–8, 2010. doi: 10.1038/nature08932.

E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.*, 10, 2009. doi: 10.1038/nrn2575.

Bureau of Public Roads. US Department of Commerce, Urban Planning Devision, Washington DC, 1964.

D.S. Callaway, M.E.J. Newman, S.H. Strogatz, and D.J. Watts. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.*, 85:5468–5471, 2000. doi: 10.1103/PhysRevLett.85.5468.

B.A. Carreras, V.E. Lynch, I. Dobson, and D.E. Newman. Critical points and transitions in an electric power transmission model for cascading failure blackouts. *Chaos*, 12(4):985–994, 2002. doi: 10.1063/1.1505810.

D. Cellai, E. López, J. Zhou, J.P. Gleeson, and G. Bianconi. Percolation in multiplex networks with overlap. *Phys. Rev. E*, 88:052811, 2013. doi: 10.1103/PhysRevE.88.052811.

Y. Chen, R. Hesse, D. Turschner, and H. Beck. Improving the grid power quality using virtual synchronous machines. In *2011 international conference on power engineering, energy and electrical drives*, pages 1–6. IEEE, 2011. doi: 10.1109/powereng.2011.6036498.

F. Cheng, I.A. Kovács, and A. Barabási. Network-based prediction of drug combinations. *Nat. Commun.*, 10(1):1197, 2019. doi: 10.1038/s41467-019-09186-x.

Z. Cheng and S. Redner. Scaling theory of fragmentation. *Phys. Rev. Lett.*, 60:2450–2453, Jun 1988. doi: 10.1103/PhysRevLett.60.2450. URL https://link.aps.org/doi/10.1103/PhysRevLett.60.2450.

H. Chiang, F. Wu, and P. Varaiya. Foundations of direct methods for power system transient stability analysis. *IEEE Trans. Circuits Syst.*, 34(2): 160–173, 1987. doi: 10.1109/tcs.1987.1086115.

S. Chung. Phase-locked loop for grid-connected three-phase power conversion systems. *IEE Proceedings-Electric Power Applications*, 147(3):213–219, 2000. doi: 10.1049/ip-epa:20000328.

G.M. Coclite, M. Garavello, and B. Piccoli. Traffic flow on a road network. *SIAM J. Math. Anal.*, 36(6):1862–1886, 2005. doi: 10.1137/s0036141004402683.

J.E. Cohen and P. Horowitz. Paradoxical behaviour of mechanical and electrical networks. *Nature*, 352:699–701, 1991. doi: 10.1038/352699a0.

R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin. Breakdown of the internet under intentional attack. *Phys. Rev. Lett.*, 86(16):3682, 2001. doi: 10.1103/physrevlett.86.3682.

R. Cohen, D. Ben-Avraham, and S. Havlin. Percolation critical exponents in scale-free networks. *Phys. Rev. E*, 66(3):036113, 2002. doi: 10.1103/

physreve.66.036113.

S. Çolak, A. Lima, and M.C. González. Understanding congested travel in urban areas. *Nat. Commun.*, 7:10793 EP, 03 2016. doi: 10.1038/ncomms10793.

R. Cole, Y. Dodis, and T. Roughgarden. How much can taxes help selfish routing? *JCSS*, 72(3):444 – 467, 2006. doi: 10.1016/j.jcss.2005.09.010.

T. Coletta and P. Jacquod. Linear stability and the braess paradox in coupled-oscillator networks and electric power grids. *Phys. Rev. E*, 93: 032222, 2016. doi: 10.1103/PhysRevE.93.032222.

S. Coombes and R. Thul. Synchrony in networks of coupled non-smooth dynamical systems: Extending the master stability function. *Eur. J. Appl. Math.*, 27(6):904–922, 2016. doi: 10.1017/s0956792516000115.

S. Coombes, Y.M. Lai, M. Şayli, and R. Thul. Networks of piecewise linear neural mass models. *Eur. J. Pure Appl.*, 29(5):869–890, 2018. doi: 10.1017/s0956792518000050.

F. Corson. Fluctuations and redundancy in optimal transport networks. *Phys. Rev. Lett.*, 104:048703, 2010. doi: 10.1103/PhysRevLett.104.048703.

P. Crucitti, V. Latora, and M. Marchiori. Model for cascading failures in complex networks. *Phys. Rev. E*, 69:045104, 2004. doi: 10.1103/PhysRevE.69.045104.

Michael F. Dacey and W. C. Krumbein. Models of breakage and selection for particle size distributions. *Journal of the International Association for Mathematical Geology*, 11(2):193–222, 1979. doi: 10.1007/BF01028965. URL https://doi.org/10.1007/BF01028965.

M. De Domenico and A. Arenas. Modeling structure and resilience of the dark network. *Phys. Rev. E*, 95:022313, 2017. doi: 10.1103/PhysRevE.95.022313.

M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M.A. Porter, S. Gómez, and A. Arenas. Mathematical formulation of multilayer

networks. *Phys. Rev. X*, 3:041022, 2013. doi: 10.1103/PhysRevX.3.041022.

E.D. Demaine, M.T. Hajiaghayi, H. Mahini, and M. Zadimoghaddam. The price of anarchy in network creation games. In *Proceedings of the Twenty-sixth Annual ACM Symposium on Principles of Distributed Computing*, pages 292–298, 2007. doi: 10.1145/1281100.1281142.

I. Dobson, B.A. Carreras, V.E. Lynch, and D.E. Newman. Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos*, 17(2):026103, 2007. doi: 10.1063/1.2737822.

F. Dörfler and F. Bullo. Synchronization in complex networks of phase oscillators: A survey. *Automatica*, 50(6):1539–1564, 2014. doi: 10.1016/j.automatica.2014.04.012.

R.M. D'Souza. Curtailing cascading failures. *Science*, 358(6365):860–861, 2017. doi: 10.1126/science.aaq0474.

B. Efron and R. Tibshirani. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statist. Sci.*, 1(1): 54–75, 1986. doi: 10.1214/ss/1177013815.

P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

E. Estrada. Covid-19 and sars-cov-2. modeling the present, looking at the future. *Phys. Rep.*, 869:1 – 51, 2020. doi: 10.1016/j.physrep.2020.07.005.

E. Estrada and N. Hatano. Communicability in complex networks. *Phys. Rev. E*, 77:036111, 2008. doi: 10.1103/PhysRevE.77.036111.

ETAP. http://bwrcs.eecs.berkeley.edu/Classes/IcBook/SPICE/.

A. Fabrikant, A. Luthra, E. Maneva, C.H. Papadimitriou, and S. Shenker. On a network creation game. In *Proceedings of the Twenty-second Annual Symposium on Principles of Distributed Computing*, pages 347–351, 2003. doi: 10.1145/872035.872088.

Facebook. Facebook data for good: Electrical distribution maps. https://dataforgood.fb.com/tools/electrical-distribution-grid-maps/, 2020. Accessed: 2020-12-01.

Y. Fang, N. Pedroni, and E. Zio. Optimization of cascade-resilient electrical infrastructures and its validation by power flow modeling. *Risk Anal.*, 35 (4):594–607, 2015. doi: 10.1111/risa.12396.

L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, and C. Fraser. Quantifying sars-cov-2 transmission suggests epidemic control with digital contact tracing. *Science*, 368(6491), 2020. doi: 10.1126/science.abb6936.

A. Fiat, H. Kaplan, M. Levy, and S. Olonetsky. Strong price of anarchy for machine load balancing. In *Automata, Languages and Programming*, pages 583–594, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg. ISBN 978-3-540-73420-8.

G. Filatrella, A.H. Nielsen, and N.F. Pedersen. Analysis of a power grid using a kuramoto-like model. *The European Physical Journal B*, 61(4):485–491, 2008. doi: 10.1140/epjb/e2008-00098-8.

A. Fouad and V. Vittal. *Power system transient stability analysis using the transient energy function method*. Pearson Education, 1991.

J. Frank and J. Massey. The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951. doi: 10.1080/01621459.1951.10500769. URL https://www.tandfonline.com/doi/abs/10.1080/01621459.1951.10500769.

D. Garlaschelli, A. Capocci, and G. Caldarelli. Self-organized network evolution coupled to extremal dynamics. *Nat. Phys.*, 3(11):813–817, 2007. doi: 10.3731/topologica.1.017.

J.D. Glover, T. Overbye, and S. Mulukutla. *Power System Analysis and Design*. CL Engineering, 2016. ISBN 130563618X.

K. Goh, D. Lee, B. Kahng, and D. Kim. Sandpile on scale-free networks.

*Phys. Rev. Lett.*, 91(14):148701, 2003. doi: 10.1103/physrevlett.91.148701.

S. Gomez, A. Diaz-Guilera, J. Gomez-Gardenes, C.J. Perez-Vicente, Y. Moreno, and A. Arenas. Diffusion dynamics on multiplex networks. *Phys. Rev. Lett.*, 110(2):028701, 2013. doi: 10.1103/physrevlett.110.028701.

J. Gómez-Gardeñes, I. Reinares, A. Arenas, and L.M. Floría. Evolution of cooperation in multiplex networks. *Scientific Reports*, 2012. doi: 10.1038/srep00620.

A.A. Hagberg, D.A. Schult, and P.J. Swart. Exploring network structure, dynamics, and function using networkx. In *Proceedings of the 7th Python in Science Conference*, pages 11 – 15, Pasadena, CA USA, 2008.

A.G. Haldane and R.M. May. Systemic risk in banking ecosystems. *Nature*, 469(7330):351, 2011. doi: 10.1038/nature09659.

A. Halu, A. Scala, A. Khiyami, and M.C. González. Data-driven modeling of solar-powered urban microgrids. *Sci. Adv.*, 2(1), 2016. doi: 10.1126/sciadv.1500700.

G. Hardin. The tragedy of the commons. *Science*, 1968. doi: 10.1126/science.162.3859.1243.

J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat. Acad. Sci. USA*, 79(8):2554–2558, 1982. doi: 10.1073/pnas.79.8.2554.

X. Huang, I. Vodenska, S. Havlin, and H.E. Stanley. Cascading failures in bi-partite graphs: Model for systemic risk propagation. *Sci. Rep.*, 3:1219, 2013. doi: 10.1038/srep01591.

F. Iannelli, A. Koher, D. Brockmann, P. Hövel, and I.M. Sokolov. Effective distances for epidemics spreading on complex networks. *Phys. Rev. E*, 95:012313, 2017. doi: 10.1103/PhysRevE.95.012313.

H. Jeong, S.P. Mason, A. Barabási, and Z.N. Oltvai. Lethality and centrality in protein networks. *Nature*, 411, 2001. doi: 10.1038/35075138.

L Katz. A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43, 1953. doi: 10.1007/BF02289026.

M. Kitsak, L.K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H.E. Stanley, and H.A. Makse. Identification of influential spreaders in complex networks. *Nat. Phys.*, 6:888 EP, 2010. doi: 10.1038/nphys1746.

V.A. Knight and P.R. Harper. Selfish routing in public services. *Eur. J. Oper. Res.*, 230(1):122 – 132, 2013. doi: 10.1016/j.ejor.2013.04.003.

Naoki Kobayashi, Kaoru Kohyama, Yo Sasaki, and Mitsugu Matsushita. Statistical laws for food fragmentation by human mastication. *Journal of the Physical Society of Japan*, 75(8):083001, 2006. doi: 10.1143/JPSJ.75.083001. URL https://doi.org/10.1143/JPSJ.75.083001.

E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *Proceedings of the 16th Annual Conference on Theoretical Aspects of Computer Science*, STACS'99, pages 404–413, Berlin, Heidelberg, 1999. Springer-Verlag. ISBN 3-540-65691-X.

N.E. Kouvaris, S. Hata, and A.D. Guilera. Pattern formation in multiplex networks. *Sci. Rep.*, 5:10840 EP, 2015. doi: 10.1038/srep10840.

Y. Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In *International symposium on mathematical problems in theoretical physics*, pages 420–422. Springer, 1975. doi: 10.1007/bfb0013365.

Y. Lai, A.E. Motter, and T. Nishikawa. *Complex Networks*, pages 299–310. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004. doi: 10.1007/978-3-540-44485-5_14.

R. Lal. Tragedy of the global commons: soil, water and air. In *Climate Change, Intercropping, Pest Control and Beneficial Microorganisms*, pages 9–11. Springer, 2009. doi: 10.1007/978-90-481-2716-0_2.

W. Li. *Risk Assessment of Power Systems*. Wiley, 2014. ISBN 978-1-118-68670-6.

A. Lima, R. Stanojevic, D. Papagiannaki, P. Rodriguez, and M.C. González. Understanding individual routing behaviour. *J. R. Soc. Interface*, 13(116), 2016. doi: 10.1098/rsif.2016.0021.

W.F. Lloyd. *Lectures on the checks to population*. Oxford University, 1833.

London Datastore. https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households, 2014. Accessed: 2019-01-20.

London Datastore. https://data.london.gov.uk/dataset/photovoltaic--pv--solar-panel-energy-generation-data, 2016. Accessed: 2019-01-21.

M. Loxton, R. Truskett, B. Scarf, L. Sindone, G. Baldry, and Y. Zhao. Consumer behaviour during crises: Preliminary research on how coronavirus has manifested consumer panic buying, herd mentality, changing discretionary spending and the role of the media in influencing behaviour. *J. Risk Fin. Man.*, 13(8):166, 2020. doi: 10.3390/jrfm13080166.

D. Manik, D. Witthaut, B. Schäfer, M. Matthiae, A. Sorge, M. Rohden, E. Katifori, and M. Timme. Supply networks: Instabilities without overload. *Eur. Phys. J. Spec. Top.*, 223(12):2527–2547, 2014. doi: 10.1140/epjst/e2014-02274-y.

B.A. Maron, L. Altucci, J. Balligand, J. Baumbach, P. Ferdinandy, S. Filetti, P. Parini, E. Petrillo, E.K. Silverman, A. Barabási, and J. Loscalzo. A global network for network medicine. *NPJ Syst. Biol. Appl.*, 6(1):29, 2020. doi: 10.1038/s41540-020-00143-9.

S.F. McWhinnie. The tragedy of the commons in international fisheries: An empirical examination. *J. Environ. Econ. Manage.*, 57(3):321–333, 2009. doi: 10.1016/j.jeem.2008.07.008.

P.J Menck, J. Heitzig, N. Marwan, and J. Kurths. How basin stability complements the linear-stability paradigm. *Nat. Phys.*, 9(2):89–92, 2013. doi: 10.1038/nphys2516.

P.J. Menck, J. Heitzig, J.and Kurths, and H.J. Schellnhuber. How dead ends undermine power grid stability. *Nat. Commun.*, 5(1):1–8, 2014. doi:

10.1038/ncomms4969.

Metro. 'my little crony' map shows web of deals between government, mps and tory donors. https://metro.co.uk/2020/11/13/my-little-crony-map-shows-deals-between-government-mps-and-tory-donors-13587900/, 2020. Accessed: 2020-11-13.

J.V. Milanović and W. Zhu. Modeling of interconnected critical infrastructure systems using complex network theory. *IEEE Trans. Smart Grid*, 9(5): 4637–4648, 2018. doi: 10.1109/TSG.2017.2665646.

I. Milchtaich. Network topology and the efficiency of equilibrium. In *ICM Millennium Lectures on Games*, pages 233–266, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg. ISBN 978-3-662-05219-8.

C. Moore and M.E.J. Newman. Epidemics and percolation in small-world networks. *Phys. Rev. E*, 61(5):5678, 2000a. doi: 10.1103/physreve.61.5678.

C. Moore and M.E.J. Newman. Exact solution of site and bond percolation on small-world networks. *Phys. Rev. E*, 62(5):7059, 2000b. doi: 10.1103/physreve.62.7059.

O. Morgenstern and J. Von Neumann. *Theory of games and economic behavior*. Princeton University Press, 1944.

A. Motter and Y. Lai. Cascade-based attacks on complex networks. *Phys. Rev. E*, 66:065102, 2003. doi: 10.1103/PhysRevE.66.065102.

A.E. Motter. Cascade control and defense in complex networks. *Phys. Rev. Lett.*, 93:098701, 2004. doi: 10.1103/PhysRevLett.93.098701.

J.F. Nash. Equilibrium points in n-person games. *Proc. Nat. Acad. Sci. USA*, 36(1):48–49, 1950. doi: 10.1073/pnas.36.1.48.

J.F. Nash. Non-cooperative games. *Ann Math*, 54(2):286–295, 1951.

R. Nicks, L. Chambon, and S. Coombes. Clusters in nonsmooth oscillator networks. *Phys. Rev. E*, 97(3):032213, 2018. doi: 10.1103/physreve.97.

032213.

T. Nishikawa and A.E. Motter. Comparative analysis of existing models for power-grid synchronization. *New J. Phys.*, 17(1):015012, 2015. doi: 10.1088/1367-2630/17/1/015012.

D.E. Olivares, A. Mehrizi-Sani, A.H. Etemadi, C.A. Cañizares, R. Iravani, M. Kazerani, A.H. Hajimiragha, O. Gomis-Bellmunt, M. Saeedifard, R. Palma-Behnke, G.A. Jiménez-Estévez, and N.D. Hatziargyriou. Trends in microgrid control. *IEEE Trans. Smart Grid*, 5(4):1905–1919, 2014. doi: 10.1109/TSG.2013.2295514.

L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, 1999. URL http://ilpubs.stanford.edu:8090/422/.

S. Pahwa, C. Scoglio, and A. Scala. Abruptness of cascade failures in power grids. *Sci. Rep.*, 4:3694, 2014. doi: 10.1038/srep03694.

S. Parhizi, H. Lotfi, A. Khodaei, and S. Bahramirad. State of the art in research on microgrids: A review. *IEEE Access*, 3:890–925, 2015. doi: 10.1109/ACCESS.2015.2443119.

R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86:3200–3203, 2001. doi: 10.1103/PhysRevLett.86.3200.

R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925–979, 2015. doi: 10.1103/RevModPhys.87.925.

G. Paul, T. Tanizawa, S. Havlin, and H.E. Stanley. Optimization of robustness of complex networks. *EPJ B*, 38(2):187–191, 2004. doi: 10.1140/epjb/e2005-00385-x.

A.C. Pigou. *The economics of welfare*. Macmillan London, 1920.

W. Quattrociocchi and A. Caldarelli, G.AND Scala. Self-healing networks:

Redundancy and structure. *PLOS ONE*, 9(2):1–7, 2014. doi: 10.1371/journal.pone.0087986.

M. Rohden, A. Sorge, M. Timme, and D. Witthaut. Self-organized synchronization in decentralized power grids. *Phys. Rev. Lett.*, 109:064101, 2012. doi: 10.1103/PhysRevLett.109.064101.

A. Rose. Efficiency in networks *Ph.D. thesis*, 2018.

A. Rose, R.D. O'Dea, and K.I. Hopcraft. Price of anarchy on heterogeneous traffic-flow networks. *Phys. Rev. E*, 94:032315, 2016. doi: 10.1103/PhysRevE.94.032315.

T. Roughgarden. The price of anarchy is independent of the network topology. *J. Comput. Syst. Sci.*, 67(2):341 – 364, 2003. doi: 10.1016/S0022-0000(03)00044-8.

T. Roughgarden. *Selfish Routing and the Price of Anarchy*. The MIT Press, 2005. ISBN 0073380571.

T. Roughgarden. On the severity of braess's paradox: Designing networks for selfish users is hard. *JCSS*, 72(5):922 – 953, 2006. doi: 10.1016/j.jcss.2005.05.009.

T. Roughgarden and É. Tardós. How bad is selfish routing? *J ACM*, 49(2):236–259, 2002. doi: 10.1145/506147.506153.

N. Schwartz, R. Cohen, D. Ben-Avraham, A. Barabási, and S. Havlin. Percolation in directed scale-free networks. *Phys. Rev. E*, 66(1):015104, 2002. doi: 10.1103/physreve.66.015104.

B. Skinner. Price of anarchy is maximized at the percolation threshold. *Phys. Rev. E*, 91:052126, 2015. doi: 10.1103/PhysRevE.91.052126.

O. Smith, J. Crowe, R.D. O'Dea, and K.I. Hopcraft. The price of anarchy in flow networks as a function of node properties. *EPL*, 127(1):18001, 2019. doi: 10.1209/0295-5075/127/18001.

O. Smith, J. Crowe, E. Farcot, R.D. O'Dea, and K.I. Hopcraft. Cascading failures in networks of heterogeneous node behavior. *Phys. Rev. E*, 101: 020301, 2020. doi: 10.1103/PhysRevE.101.020301.

A. Solé-Ribalta, S. Gómez, and A. Arenas. Congestion induced by the structure of multiplex networks. *Phys. Rev. Lett.*, 116:108701, 2016. doi: 10.1103/PhysRevLett.116.108701.

SPICE. https://etap.com/.

D. Stauffer and A. Aharony. *Introduction to percolation theory*. CRC press, 2018.

B. Steinegger, A. Arenas, J. Gómez-Gardeñes, and C. Granell. Pulsating campaigns of human prophylaxis driven by risk perception palliate oscillations of direct contact transmitted diseases. *Phys. Rev. Research*, 2:023181, 2020. doi: 10.1103/PhysRevResearch.2.023181.

B. Stott, J. Jardim, and O. Alsac. Dc power flow revisited. *IEEE Trans. Power Systems*, 24(3):1290–1300, 2009. doi: 10.1109/TPWRS.2009.2021235.

S. Tadelis. *Game theory: an introduction*. Princeton University Press, 2013.

W. Tan, M. B. Blake, I. Saleh, and S. Dustdar. Social-network-sourced big data analytics. *IEEE Internet Comput.*, 17(5):62–69, 2013. doi: 10.1109/mic.2013.100.

T. Tanizawa, G. Paul, R. Cohen, S. Havlin, and H.E. Stanley. Optimization of network robustness to waves of targeted and random attacks. *Phys. Rev. E*, 71(4):047101, 2005. doi: 10.1103/physreve.71.047101.

Tesla. Tesla powerwall 2 datasheet. https://www.tesla.com/sites/default/files/pdfs/powerwall/Powerwall%202_AC_Datasheet_en_GB.pdf, 2018. Accessed: 2020-10-10.

The Guardian. 'chumocracy': how covid revealed the new shape of the tory establishment. https://www.theguardian.com/world/2020/nov/15/chumocracy-covid-revealed-shape-tory-establishment, 2020a. Accessed:

2020-11-15.

The Guardian. How coronavirus spreads through a population and how we can beat it. https://www.theguardian.com/world/datablog/ng-interactive/2020/apr/22/see-how-coronavirus-can-spread-through-a-population-and-how-countries-flatten-the-curve, 2020b. Accessed: 2020-11-15.

The New York Times. How the virus got out. https://www.nytimes.com/interactive/2020/03/22/world/coronavirus-spread.html, 2020. Accessed: 2020-11-15.

timeanddate.com. https://www.timeanddate.com/sun, 2021. Accessed: 2021-05-08.

J.C. Vasquez, J.M. Guerrero, A. Luna, P. Rodriguez, and R. Teodorescu. Adaptive droop control applied to voltage-source inverters operating in grid-connected and islanded modes. *IEEE Trans. Ind. Electron.*, 56(10): 4088–4096, 2009. doi: 10.1109/tie.2009.2027921.

A. Vespignani, H. Tian, C. Dye, J.O. Lloyd-Smith, R.M. Eggo, M. Shrestha, S.V. Scarpino, B. Gutierrez, M.U.G. Kraemer, J. Wu, K. Leung, and G.M. Leung. Modelling covid-19. *Nat. Rev. Phys.*, 2(6):279–281, 2020. doi: 10.1038/s42254-020-0178-4.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

S. Wang and J. Baillieul. Kirchhoff-braess phenomena in dc electric networks. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 3286–3293, 2016. doi: 10.1109/CDC.2016.7798763.

D.J. Watts. A simple model of global cascades on random networks. *Proc. Nat. Acad. Sci. USA*, 99(9):5766–5771, 2002. doi: 10.1073/pnas.082090499.

D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440, 1998. doi: 10.1038/30918.

D. Witthaut and M. Timme. Braess's paradox in oscillator networks, desynchronization and power outage. *New J. Phys.*, 14(8):083036, 2012. doi: 10.1088/1367-2630/14/8/083036.

D. Witthaut and M. Timme. Nonlocal effects and countermeasures in cascading failures. *Phys. Rev. E*, 92:032809, 2015. doi: 10.1103/PhysRevE.92.032809.

Y. Yang, T. Nishikawa, and Adilson E. Motter. Small vulnerable sets determine large network cascades in power grids. *Science*, 358(6365), 2017. doi: 10.1126/science.aan3184.

H. Youn, M.T. Gastner, and H. Jeong. Price of anarchy in transportation networks: Efficiency and optimality control. *Phys. Rev. Lett.*, 101:128701, 2008. doi: 10.1103/PhysRevLett.101.128701.

Y. Zhang, A. Arenas, and O. Yağan. Cascading failures in interdependent systems under a flow redistribution model. *Phys. Rev. E*, 97:022307, 2018. doi: 10.1103/PhysRevE.97.022307.